REMARKS

Claims 1, 9-12, 14-15, 28, and 47-48 constitute the pending claims in the present application. Applicants respectfully request reconsideration in view of the following remarks. Issues raised by the Examiner will be addressed below in the order they appear in the prior Office Action.

The Claims Comply with 35 U.S.C. §112

A. Written Description

Claims 1, 9-15, 28, and 43-48 are rejected under 35 U.S.C. §112, first paragraph, as allegedly containing subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventors, at the time the application was filed, had possession of the claimed invention. Applicants respectfully traverse this rejection.

In particular, the Examiner states that neither the support cited in the previous submission nor the specification support a "*method for attenuating expression of a target gene in mammalian cells comprising introducing into mammalian cells suspended in culture an expression vector encoding a hairpin RNA*" having the limitations relating to expression from a vector, length, or avoidance of protein kinase activation as cited in the claims. Applicants respectfully submit that the specification is replete with teachings that support the currently pending claims. In addition to the support cited in the previous submission, the Examiner's attention is directed to the following passages as exemplary support: page 2, lines 20-24, page 3, lines 21-36, page 5, lines 18-22, page 13, lines 16-27, and page 14, lines 7-10. Additionally, Applicants respectfully wish to point out that the teachings of the application must be evaluated as a whole and that individual sections cannot be read in isolation.

As discussed below, as well as in the attached Declaration of Frank McKeon (herein the McKeon Declaration), various terms in the application have been defined and these definitions are applicable throughout the application whenever these terms are used. For instance, the application uses the terms "hairpin RNA" and "single self-complementary RNA", which as the McKeon Declaration avers, would be understood by a molecular biologist to be interchangeable terms referring to hairpin RNAs. Applicants respectfully contend that one of ordinary skill in the art would recognize that the instant application includes hairpin RNA species within the scope of the general term "double

stranded RNA" (or "dsRNA"). This point is also made by the McKeon Declaration. To illustrate, referring to the published application, paragraph No. 0037 states "[t]he dsRNA construct may comprise one or more strands of polymerized ribonucleotide". That paragraph goes on to state that "[t]he double-stranded structure may be formed by a single self-complementary RNA strand or two complementary RNA strands". Likewise, paragraph No. 0135 provides that the "[t]he double-stranded structure may be formed by a single self-complementary RNA strand or two complementary RNA strands".

The instant application also clearly contemplates the use of expression systems that are intended to produce hairpin RNAs upon being transcribed in cells. As the McKeon Declaration points out, exemplary teachings of these embodiments include, to illustrate, paragraph No. 0019 which states "[y]et another aspect of the present invention provides a method for attenuating expression of a target gene in cultured cells, comprising introducing an expression vector having a "coding sequence" which, when transcribed, produces double stranded RNA (dsRNA) the cell in an amount sufficient to attenuate expression of the target gene, wherein the dsRNA comprises a nucleotide sequence that hybridizes under stringent conditions to a nucleotide sequence of the target gene". Other examples of support include the statement at paragraph No. 0120 that "[t]he dsRNA construct may be synthesized either in vivo or in vitro... For generating double stranded transcripts from a transgene in vivo, a regulatory region may be used to transcribe the RNA strand (or strands)". Likewise, paragraph No. 0138 states "[t]he dsRNA construct may be synthesized either in vivo or in vitro....For transcription from a transgene in vivo or an expression construct, a regulatory region (e.g., promoter, enhancer, silencer, splice donor and acceptor, polyadenylation) may be used to transcribe the dsRNA strand (or strands)"; and paragraph No. 0139 states "[a] viral construct packaged into a viral particle would accomplish both efficient introduction of an expression construct into the cell and transcription of dsRNA construct encoded by the expression construct". Paragraph 0120 also states that "Endogenous RNA polymerase of the cell may mediate transcription in vivo, or cloned RNA polymerase can be used for transcription in vivo or in vitro".

In view of the definition of double-stranded RNA, i.e., that it includes hairpin RNAs, a molecular biologist would conclude, as asserted in the McKeon Declaration, from these and other teachings provided by the Hammond Application that the inventors had intended to convey their invention to include the use of expression constructs to produce, by transcription, hairpin RNA in cells. Moreover, working examples of these embodiments are explicitly described in the Figures and Exemplification, such as Figure

27 and Example 3 (titled "A simplified Method for the Creation of Hairpin Constructs for RNA Interference").

The present application explicitly supports gene silencing in mammalian cells using the transcribed hairpin RNAs. For example, the Hammond Application provides that the disclosed methods for gene silencing are broadly applicable to a variety of cell types including mammalian cells such as primate and human cells. Additionally, the Hammond Application teaches various expression vectors, several of which are suitable for use in mammalian cells, for generating dsRNA species in cells.

Furthermore, the instant application provides the necessary understanding of the mechanism by which RNA interference ("RNAi") works so that, in view of it's teachings and the state-of-the-art in RNA interference, one skilled in the art would appreciate that hairpin RNAs could be used to induce gene silencing in mammalian cells. As detailed in the McKeon Declaration, the present application provides the essential features of RNA interference such that, for the first time, one of ordinary skill in the art would appreciate that it was a generalized phenomenon and that it was conserved between mammalian and drosophila cells and could be used in cells in culture.

As discussed in the McKeon Declaration, at the time of the earliest priority date of the present application, i.e., March 2000, procedures based on double stranded RNA-triggered silencing were fairly well-established tools for functional genomics of lower intact (whole) organisms, such as plants, invertebrates and fungi. There were also indications in the literature that double-stranded RNA had some ability to induce gene silencing in mammals. For instance, the Wianny et al. reference, [(2000) Nature Cell Biol 2:70] showed that double-stranded RNA, in this case in the form of a two stranded duplex, could effect gene function in early mouse embryos.

As recognized and discussed by the McKeon Declaration, the biochemical and genetic data provided by the teachings of the present application provided an adequate understanding of the mechanisms underlying RNA interference in both invertebrates and vertebrates such that one skilled in the art would have readily believed that hairpin RNAs could be used for inducing gene silencing in mammalian cells. In particular, the present application identifies the existence of conserved machinery for double stranded RNA-induced gene silencing between drosophila to mammals. The application also defines the RNA interference pathway as proceeding via a two-step mechanism, and showed that the fundamental enzymes involved in that processes, namely the RNase III nuclease called

Dicer and the key protein in the RISC complex Argonaute, were conserved between drosophila and mammals, and therefore provides the necessary guidance for one skilled in the art to appreciate that the components of the RNA interference pathway are conserved between drosophila and mammals. Further, Example 2 of the application makes several significant points. First is that the dicer enzyme was the nuclease responsible for the processing of double-stranded RNA to form guide sequences of about 22 nucleotides. Second is that the human dicer homolog was also capable of generating guide sequences of about 22 nucleotides. See also Figure 23.

The McKeon Declaration states that the totality of these teachings, and the state-of-the-art, led him to conclude from the application as filed that the inventors had possession of the invention of the use of hairpin RNA to induce gene silencing in mammalian cells. To summarize, the basis of Dr. McKeon's belief includes, inter alia, teachings from the instant application of:

- the elucidation of role of a nuclease activity, dicer, that cleaves regions of double-stranded RNA;
- the conservation of dicer between drosophila and mammals;
- the conservation of the dicer product, e.g., the ~22-mer products between drosophila and mammals; and
- the ability of hairpin RNA to induce gene silencing.

The McKeon Declaration also noted the state-of-the-art in using other double stranded RNA species, and that one skilled in the art would have concluded from the teachings of present application that these would also be substrates for dicer. For these reasons, it is submitted that a molecular biologist reading the present application would clearly conclude that the inventors had possession of, and had clearly described, the use of transcribed hairpin RNAs for gene silencing in mammalian cells. That is, the Hammond Application provides a clear written description of the use of hairpin RNAs for gene silencing in mammalian cells.

With respect to claim 13 (which has been canceled), Applicants aver that in view of the above arguments, one of ordinary skill in the art would have reasonably implied from the specification of the present that the inventors contemplated embodiments in which the double stranded portion of the hairpin RNA was at least 20 nucleotides in length. For instance, the specification teaches that the double stranded portion of the hairpin is a substrate for dicer, and also states that the "dsRNA is at least 20 nucleotides

in length". However, in order to expedite prosecution of the instant application, claim 13 has been canceled.

It is respectfully submitted that maintenance of the present rejection is not supported by substantial evidence, and as such, does not meet the "arbitrary, capricious" standard applied under the "substantial evidence" test of Section 706(2)(E) of the Administrative Procedure Act. Dickinson v. Zurko, 119 S. Ct. 1816 (1999). Indeed, the CAFC has stated "deferential review under the [APA] does not relieve the agency of its obligation to develop an evidentiary basis for its findings. To the contrary, the [APA] reinforces this obligation." In re Sang Su Lee Case Number 00-1158 (CAFC January 18, 2002). No relevant or other fact finding has been cited or relied upon by the Examiner to rebut the Applicants' arguments.

B. Enablement

Claims 1, 9-15, 28, and 43-48 are rejected under 35 U.S.C. §112, first paragraph, as allegedly containing subject matter which was not described in the specification in such a way as to enable one skilled in the art to which it pertains, or with which it is most nearly connected, to make and/or use the invention. Applicants respectfully traverse this rejection.

The Examiner states that "[t]he specification as filed only provides sufficient guidance and/or instruction for using double stranded RNA within cells, specifically Drosophila cells, in an *in vitro* environment, but provides no support anywhere for using 'hairpin RNA' or 'transcribed hairpin RNA' ... in an *in vitro* environment or in an *in vivo* environment, which also encompass cells in a whole organisms." The Examiner further states that the specification does not provide sufficient guidance for a method of "gene therapy" and that there is no evidence that "enough mammalian cells can be transfected to provide any therapeutic benefits." Applicants respectfully point out the currently pending claims are directed to methods for attenuating expression of a target gene in mammalian cells suspended in culture, e.g., isolated from the whole organism and provided in culture media. Accordingly, the claims do not encompass transfection of cells in a whole organism or methods of gene therapy.

Applicants note that "[c]ompliance with the enablement requirement of 35 U.S.C. § 112, first paragraph, does not turn on whether an example is disclosed. An example may be 'working' or 'prophetic.'" Further, "[t]he specification need not contain an

example if the invention is otherwise disclosed in such a manner that one skilled in the art will be able to practice it without an undue amount experimentation. *In re Borkowski*, 442, F.2d 904, 908, 164 USPQ 642,645 (CCPA 1970)." See MPEP 2164.02.

With regard to attenuation of gene expression in mammalian cells suspended in culture, Applicants submit that the specification provides sufficient guidance for one skilled in the art to make and/or use the claimed invention without undue experimentation. In particular, the specification provides working examples of attenuation of gene expression not only in Drosophila S2 cells but also in mammalian cells (see, e.g., page 10, line 24 to page 11, line 3 and page 11, lines 26-35). In fact, at least human, hamster, and mouse cell lines were shown to be effective mammalian cell targets for gene attenuation purposes (see, e.g., Figures 28-33). Additionally, the specification teaches a method for creating hairpin constructs for RNA interference and states that such constructs have been successfully used to provoke gene silencing in drosophila cells (see e.g., Example 3 at page 43, line 24 to page 44, line 14). The specification further teaches a variety of expression vectors for use in association with mammalian cells (see e.g., page 3, line 31 to page 4, line 9 and page 19, lines 4-18). Given the ample teachings of the specification and the knowledge in the art at the time the application was filed, one skilled in the art could readily practice gene attenuation in mammalian cells using transcribed hairpin RNA without undue experimentation.

Additionally, Applicants provide herewith the Declaration of Gregory Hannon (herein the Hannon Declaration) and certain scientific papers attached thereto as Exhibits A-D. Exhibits A and B are papers subsequently published by Applicants that disclose successful gene attenuation in a variety of mammalian cell lines using transcribed hairpin RNAs. Exhibits C and D describe work by Applicants in which libraries of expression vectors for production of hairpin RNAs in mammalian cells were successfully produced and tested. The papers provides evidence that the teachings of the specification broadly enable gene silencing using transcribed hairpin RNAs in mammalian cells suspended in culture. Also, the Hannon Declaration sets forth evidence of the ease with which transcribed hairpin RNAs can be used to attenuate gene expression in a wide variety of mammalian cell lines using the teachings of the specification.

The McKeon Declaration also provides further support that Applicants' teachings fully enable one of ordinary skill to make and use transcribed hairpin RNAs for gene silencing in mammalian cell lines. As Dr. McKeon states, no more than routine experimentation is required to select a target sequence in a mammalian gene to be silenced, design the hairpin motif that should include that sequence, generate an

expression vector for transcribing the hairpin in a mammalian cell, and assess whether or not the vector results in down-regulation of the level of expression of that gene, such as by reduced mRNA levels. Each step of this process involves routine molecular biology. As the McKeon Declaration attests, hairpin RNA constructs are routinely designed and generated by graduate students in his laboratory at Harvard Medical School.

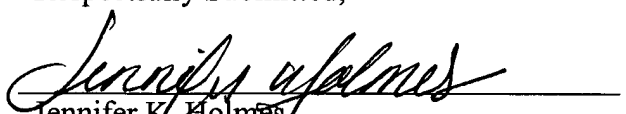
Applicants note that the information and statements set forth in the attached declarations, though based in part on work carried out after Applicants' filing date, can be used to substantiate the credibility of the asserted enablement of the subject invention with respect to other embodiments within the scope of the claims. This evidence is not offered to render an insufficient disclosure enabling, rather, they prove that the disclosure was in fact enabling when filed. In re Brana, 51 F2d 1560 (CAFC 1995).

For the reasons presented above, Applicants submit that the claims fully comply with the enablement requirement under 35 U.S.C. § 112, first paragraph. Reconsideration and withdrawal of this rejection are respectfully requested.

CONCLUSION

In view of the foregoing amendments and remarks, Applicants submit that the pending claims are in condition for allowance. Early and favorable reconsideration is respectfully solicited. The Examiner may address any questions raised by this submission to the undersigned at 617-951-7000. Should any additional extension of time be required, Applicants hereby petition for same and request that the extension fee and any other fee required for timely consideration of this submission be charged to **Deposit Account No. 18-1945.**

Respectfully Submitted,



Jennifer K. Holmes
Reg. No. 46,778

Date: August 9, 2005

Customer No: 28120
Docketing Specialist
Ropes & Gray LLP
One International Place
Boston, MA 02110
Phone: 617-951-7739
Fax: 617-951-7050

Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells

Patrick J. Paddison,[1] Amy A. Caudy,[1] Emily Bernstein,[2,3] Gregory J. Hannon,[1,2,4] and Douglas S. Conklin[2]

[1]Watson School of Biological Sciences, [2]Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA; [3]Graduate Program in Genetics, State University of New York at Stony Brook, Stony Brook, New York 11794, USA

RNA interference (RNAi) was first recognized in *Caenorhabditis elegans* as a biological response to exogenous double-stranded RNA (dsRNA), which induces sequence-specific gene silencing. RNAi represents a conserved regulatory motif, which is present in a wide range of eukaryotic organisms. Recently, we and others have shown that endogenously encoded triggers of gene silencing act through elements of the RNAi machinery to regulate the expression of protein-coding genes. These small temporal RNAs (stRNAs) are transcribed as short hairpin precursors (~70 nt), processed into active, 21-nt RNAs by Dicer, and recognize target mRNAs via base-pairing interactions. Here, we show that short hairpin RNAs (shRNAs) can be engineered to suppress the expression of desired genes in cultured *Drosophila* and mammalian cells. shRNAs can be synthesized exogenously or can be transcribed from RNA polymerase III promoters in vivo, thus permitting the construction of continuous cell lines or transgenic animals in which RNAi enforces stable and heritable gene silencing.

[*Key Words*: RNAi; gene silencing; miRNA; shRNA; siRNA]

Received January 31, 2002; revised version accepted March 8, 2002.

An understanding of the biological role of any gene comes only after observing the phenotypic consequences of altering the function of that gene in a living cell or organism. In many cases, those organisms for which convenient methodologies for genetic manipulation exist blaze the trail toward an understanding of similar genes in less tractable organisms, such as mammals. The advent of RNA interference (RNAi) as an investigational tool has shown the potential to democratize at least one aspect of genetic manipulation, the creation of hypomorphic alleles, in organisms ranging from unicellular parasites (e.g., Shi et al. 2000) to mammals (Svoboda et al. 2000; Wianny and Zernicka-Goetz 2000).

Although *Caenorhabditis elegans* has, for some time, been well developed as a forward genetic system, the lack of methodologies for gene replacement by homologous recombination presented a barrier to assessing rapidly the consequences of loss of function in known genes. In an effort to overcome this limitation, Mello and Fire (Fire et al. 1998), building on earlier studies (Guo and Kemphues 1995), probed the utility of antisense RNA as a method for suppressing gene expression in worms. Through these efforts, they found that double-stranded RNA (dsRNA) was much more effective than antisense RNA as an inducer of gene silencing. Subsequent studies have shown that RNAi is a conserved biological response that is present in many, if not most, eukaryotic organisms (for review, see Bernstein et al. 2001b; Hammond et al. 2001b).

As a result of biochemical and genetic approaches in several experimental systems, the mechanisms underlying RNAi have begun to unfold (for review, see Bernstein et al. 2001b; Hammond et al. 2001b). These suggest the existence of a conserved machinery for dsRNA-induced gene silencing, which proceeds via a two-step mechanism. In the first step, the dsRNA silencing trigger is recognized by an RNase III family nuclease called Dicer, which cleaves the dsRNA into ~21–23-nt siRNAs (small interfering RNAs). These siRNAs are incorporated into a multicomponent nuclease complex, RISC, which identifies substrates through their homology to siRNAs and targets these cognate mRNAs for destruction.

Although it was clear from the outset that RNAi would prove a powerful tool for manipulating gene expression in invertebrates, there were several potential impediments to the use of this approach in mammalian cells. Most mammalian cells harbor a potent antiviral response that is triggered by the presence of dsRNA viral

[4]Corresponding author.
E-MAIL hannon@cshl.org; FAX (516) 367-8874.
Article and publication are at http://www.genesdev.org/cgi/doi/10.1101/gad.981002.

replication intermediates. A key component of this response is a dsRNA-activated protein kinase, PKR, which phosphorylates EIF-2α, inducing, in turn, a generalized inhibition of translation (for review, see Williams 1997; Gil and Esteban 2000). In addition, dsRNA activates the 2′5′ oligoadenylate polymerase/RNase L system and represses IκB. The ultimate outcome of this set of responses is cell death via apoptosis.

Therefore, it came as a welcome surprise that dsRNA could induce sequence-specific silencing in mammalian embryos, which apparently lack generalized responses to dsRNA (Svoboda et al. 2000; Wianny and Zernicka-Goetz 2000). Indeed, microinjection of dsRNA into mouse zygotes could specifically silence both exogenous reporters and endogenous genes to create anticipated phenotypes. Subsequently, these observations were extended to embryonic cell lines, such as embryonic stem cells and embryonal carcinoma cells, which do not show generic translational repression in response to dsRNA (Billy et al. 2001; Yang et al. 2001; Paddison et al. 2002). However, restriction of conventional RNAi to these few embryonic and cell culture systems would place a significant limitation on the utility of this approach in mammals.

Tuschl and colleagues first showed that short RNA duplexes, designed to mimic the products of the Dicer enzyme, could trigger RNA interference in vitro in *Drosophila* embryo extracts (Tuschl et al. 1999; Elbashir et al. 2001b,c). This observation was extended to mammalian somatic cells by Tuschl and coworkers (Elbashir et al. 2001a) and by Fire and colleagues (Caplen et al. 2001), who showed that chemically synthesized siRNAs could induce gene silencing in a wide range of human and mouse cell lines. The use of synthetic siRNAs to transiently suppress the expression of target genes is quickly becoming a method of choice for probing gene function in mammalian cells.

Dicer, the enzyme that normally produces siRNAs in vivo, has been linked to RNA interference both through biochemistry and through genetics (Bernstein et al. 2001a; Grishok et al. 2001; Ketting et al. 2001; Knight and Bass 2001). Indeed, *C. elegans* animals that lack Dicer are RNAi-deficient, at least in some tissues. However, these animals also have additional phenotypic abnormalities. Specifically, they are sterile and show a number of developmental abnormalities that typify alterations in developmental timing. Indeed, the phenotypes of the *Dicer* mutant animals were similar to those previously observed for animals carrying mutations in the *let-7* gene (Reinhart et al. 2000).

The *let-7* gene encodes a small, highly conserved RNA species that regulates the expression of endogenous protein-coding genes during worm development. The active RNA species is transcribed initially as an ~70-nt precursor, which is posttranscriptionally processed into a mature ~21-nt form (Reinhart et al. 2000). Both in vitro and in vivo data from *C. elegans* (Grishok et al. 2001; Ketting et al. 2001; Knight and Bass 2001) and human cells (Hutvagner et al. 2001) have pointed to Dicer as the enzyme responsible for *let-7* maturation and for the maturation of a similar small RNA, *lin-4* (Grishok et al. 2001). Thus, at least some components of the RNAi machinery respond to endogenously encoded triggers to regulate the expression of target genes.

Recent studies have placed *let-7* and *lin-4* as the founding members of a potentially very large group of small RNAs known generically as micro-RNAs (miRNAs). Nearly 100 potential miRNAs have now been identified in *Drosophila*, *C. elegans*, and mammals (Lagos-Quintana et al. 2001; Lau et al. 2001; Lee and Ambros 2001). Although the functions of these diverse RNAs remain mysterious, it seems likely that they, like *let-7* and *lin-4*, are transcribed as hairpin RNA precursors, which are processed to their mature forms by Dicer (Lee and Ambros 2001; E. Bernstein, unpubl.).

Since the realization that small, endogenously encoded hairpin RNAs could regulate gene expression via elements of the RNAi machinery, we have sought to exploit this biological mechanism for the regulation of desired target genes. Here we show that short hairpin RNAs (shRNAs) can induce sequence-specific gene silencing in mammalian cells. As is normally done with siRNAs, silencing can be provoked by transfecting exogenously synthesized hairpins into cells. However, silencing can also be triggered by endogenous expression of shRNAs. This observation opens the door to the production of continuous cells lines in which RNAi is used to stably suppress gene expression in mammalian cells. Furthermore, similar approaches should prove efficacious in the creation of transgenic animals and potentially in therapeutic strategies in which long-term suppression of gene function is essential to produce a desired effect.

Results

Short hairpin RNAs trigger gene silencing in Drosophila *cells*

Several groups (Grishok et al. 2001; Hutvagner et al. 2001; Ketting et al. 2001; Knight and Bass 2001) have shown that endogenous triggers of gene silencing, specifically small temporal RNAs (stRNAs) *let-7* and *lin-4*, function at least in part through RNAi pathways. Specifically, these small RNAs are encoded by hairpin precursors that are processed by Dicer into mature, ~21-nt forms. Moreover, genetic studies in *C. elegans* have shown a requirement for Argonaute-family proteins in stRNA function. Specifically, alg-1 and alg-2, members of the EIF2c subfamily, are implicated both in stRNA processing and in their downstream effector functions (Grishok et al. 2001). We have recently shown that a component of RISC, the effector nuclease of RNAi, is a member of the Argonaute family, prompting a model in which stRNAs may function through RISC-like complexes, which regulate mRNA translation rather than mRNA stability (Hammond et al. 2001a).

We wished to test the possibility that we might retarget these small, endogenously encoded hairpin RNAs to regulate genes of choice with the ultimate goal of sub-

verting this regulatory system for manipulating gene expression stably in mammalian cell lines and in transgenic animals. Whether triggered by long dsRNAs or by siRNAs, RNAi is generally more potent in the suppression of gene expression in *Drosophila* S2 cells than in mammalian cells. We therefore chose this model system in which to test the efficacy of short hairpin RNAs (shRNAs) as inducers of gene silencing.

Neither stRNAs nor the broader group of miRNAs that has recently been discovered form perfect hairpin structures. Indeed, each of these RNAs is predicted to contain several bulged nucleotides within their rather short (~30-nt) stem structures. Because the position and character of these bulged nucleotides have been conserved throughout evolution and among at least a subset of miRNAs, we sought to design retargeted miRNA mimics to conserve these predicted structural features. Only the *let-7* and *lin-4* miRNAs have known mRNA targets (Wightman et al. 1993; Slack et al. 2000). In both cases, pairing to binding sites within the regulated transcripts is imperfect, and in the case of *lin-4*, the presence of a bulged nucleotide is critical to suppression (Ha et al. 1996). We therefore also designed shRNAs that paired imperfectly with their target substrates. A subset of these shRNAs is depicted in Figure 1A.

To permit rapid testing of large numbers of shRNA variants and quantitative comparison of the efficacy of suppression, we chose to use a dual-luciferase reporter system, as previously described for assays of RNAi in both *Drosophila* extracts (Tuschl et al. 1999) and mammalian cells (Caplen et al. 2001; Elbashir et al. 2001a). Cotransfection of firefly and *Renilla* luciferase reporter plasmids with either long dsRNAs or with siRNAs homologous to the firefly luciferase gene yielded an ~95% suppression of firefly luciferase without effect on *Renilla* luciferase (Fig. 1B; data not shown). Firefly luciferase could also be specifically silenced by cotransfection with homologous shRNAs. Surprisingly, those shRNAs modeled most closely on the *let-7* paradigm were the least effective inducers of silencing (data not shown). The inclusion of bulged nucleotides within the shRNA stem caused only a modest reduction in potency; however, the presence of mismatches with respect to the target mRNA essentially abolished silencing potential. The most potent inhibitors were those composed of simple hairpin structures with complete homology to the sub-

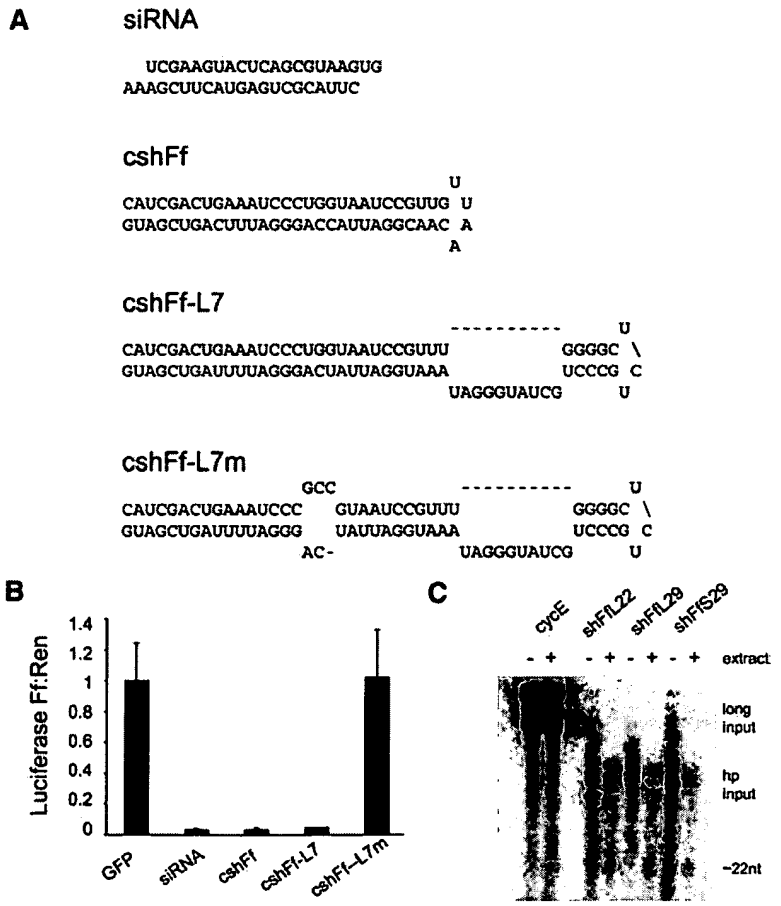


Figure 1. Short hairpins suppress gene expression in *Drosophila* S2 cells. (*A*) Sequences and predicted secondary structure of representative chemically synthesized RNAs. Sequences correspond to positions 112–134 (siRNA) and 463–491 (shRNAs) of Firefly luciferase carried on pGL3-Control. An siRNA targeted to position 463–485 of the luciferase sequence was virtually identical to the 112–134 siRNA in suppressing expression, but is not shown. (*B*) Exogenously supplied short hairpins suppress expression of the targeted Firefly luciferase gene in vivo. Six-well plates of S2 cells were transfected with 250 ng/well of plasmids that direct the expression of firefly and *Renilla* luciferase and 500 ng/well of the indicated RNA. Luciferase activities were assayed 48 h after transfection. Ratios of firefly to *Renilla* luciferase activity were normalized to a control transfected with an siRNA directed at the green fluorescent protein (GFP). The average of three independent experiments is shown; error bars indicate standard deviation. (*C*) Short hairpins are processed by the *Drosophila* Dicer enzyme. T7 transcribed hairpins shFfL22, shFfL29, and shFfS29 were incubated with (+) and without (–) 0–2-h *Drosophila* embryo extracts. Those incubated with extract produced ~22-nt siRNAs, consistent with the ability of these hairpins to induce RNA interference. A long dsRNA input (cyclin E 500-mer) was used as a control. Cleavage reactions were performed as described in Bernstein et al. (2001a).

strate. Introduction of G-U basepairs either within the stem or within the substrate recognition sequence had little or no effect (Fig. 1A,B; data not shown). Similarly, varying either the loop size from ~4 to 23 bases or the loop sequence (e.g., to mimic *let-7*) also proved neutral (data not shown).

These results show that short hairpin RNAs can induce gene silencing in *Drosophila* S2 cells with potency similar to that of siRNAs (Fig. 1B). However, in our initial observation of RNA interference in *Drosophila* S2 cells, we noted a profound dependence of the efficiency of silencing on the length of the dsRNA trigger (Hammond et al. 2000). Indeed, dsRNAs of fewer than ~200 nt triggered silencing very inefficiently. Silencing is initiated by an RNase III family nuclease, Dicer, that processes long dsRNAs into ~22-nt siRNAs. In accord with their varying potency as initiators of silencing, long dsRNAs are processed much more readily than short RNAs by the Dicer enzyme (Bernstein et al. 2001a). We therefore tested whether shRNAs were substrates for the Dicer enzyme.

We had noted previously that *let-7* (Ketting et al. 2001) and other miRNAs (E. Bernstein, unpubl.) are processed by Dicer with an unexpectedly high efficiency as compared with short, nonhairpin dsRNAs. Similarly, Dicer efficiently processed shRNAs that targeted firefly luciferase, irrespective of whether they were designed to mimic a natural Dicer substrate (*let-7*) or whether they were simple hairpin structures (Fig. 1C). These data suggest that recombinant shRNAs can be processed by Dicer into siRNAs and are consistent with the idea that these short hairpins trigger gene silencing via an RNAi pathway.

Short hairpin activated gene silencing in mammalian cells

RNAi is developing into an increasingly powerful methodology for manipulating gene expression in diverse experimental systems. However, mammalian cells contain several endogenous systems that were predicted to hamper the application of RNAi. Chief among these is a dsRNA-activated protein kinase, PKR, which effects a general suppression of translation via phosphorylation of EIF-2α (Williams 1997; Gil and Esteban 2000). Activation of these, and other dsRNA-responsive pathways, generally requires duplexes exceeding 30 bp in length, possibly to permit dimerization of the enzyme on its allosteric activator (e.g., Clarke and Mathews 1995).

Small RNAs that mimic Dicer products, siRNAs, presumably escape this limit and trigger specific silencing, in part because of their size. However, short duplex RNAs that lack signature features of siRNAs can efficiently induce silencing in *Drosophila* S2 cells but not in mammalian cells (A.A. Caudy, unpubl.). Endogenously encoded miRNAs may also escape PKR surveillance because of their size but perhaps also because of the discontinuity of their duplex structure. Given that shRNAs of <30 bp were effective inducers of RNAi in *Drosophila* S2 cells, we tested whether these RNAs could also induce sequence-specific silencing in mammalian cells.

Human embryonic kidney (HEK293T) cells were cotransfected with chemically synthesized shRNAs and with a mixture of firefly and *Renilla* luciferase reporter plasmids. As had been observed in S2 cells, shRNAs were effective inducers of gene silencing. Once again, hairpins designed to mimic *let-7* were consistently less effective than were simple hairpin RNAs, and the introduction of mismatches between the antisense strand of the shRNA and the mRNA target abolished silencing (Fig. 2A; data not shown). Overall, shRNAs were somewhat less potent silencing triggers than were siRNAs. Whereas siRNAs homologous to firefly luciferase routinely yielded ~90%–95% suppression of gene expression, suppression levels achieved with shRNAs ranged from 80%–90% on average. As we also observe with siRNAs, the most important determinant of the potency of the silencing trigger is its sequence. We find that roughly 50% of both siRNAs and shRNAs are competent for suppressing gene expression. However, neither analysis of the predicted structures of the target mRNA nor analysis of alternative structures in siRNA duplexes or shRNA hairpins has proved of predictive value for choosing effective inhibitors of gene expression.

We have adopted as a standard, shRNA duplexes containing 29 bp. However, the size of the helix can be reduced to ~25 nt without significant loss of potency. Duplexes as short as 22 bp can still provoke detectable silencing, but do so less efficiently than do longer duplexes. In no case do we observe a reduction in the internal control reporter (*Renilla* luciferase) that would be consistent with an induction of nonspecific dsRNA responses.

The ability of shRNAs to induce gene silencing was not confined to 293T cells. Similar results were also obtained in a variety of other mammalian cell lines, including human cancer cells (HeLa), transformed monkey epithelial cells (COS-1), murine fibroblasts (NIH 3T3), and diploid human fibroblasts (IMR90; Fig. 2; data not shown).

Synthesis of effective inhibitors of gene expression using T7 RNA polymerase

The use of siRNAs to provoke gene silencing is developing into a standard methodology for investigating gene function in mammalian cells. To date, siRNAs have been produced exclusively by chemical synthesis (e.g., Caplen et al. 2001; Elbashir et al. 2001a). However, the costs associated with this approach are significant, limiting its potential utility as a tool for investigating in parallel the functions of large numbers of genes. Short hairpin RNAs are presumably processed into active siRNAs in vivo by Dicer (see Fig. 1C). Thus, these may be more tolerant of terminal structures, both with respect to nucleotide overhangs and with respect to phosphate termini. We therefore tested whether shRNAs could be prepared by in vitro transcription with T7 RNA polymerase.

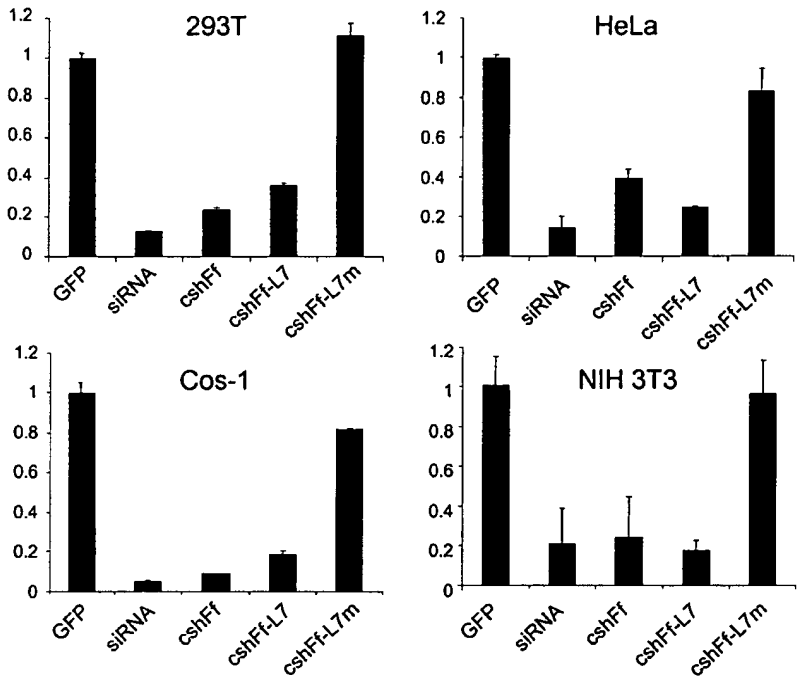


Figure 2. Short hairpins function in mammalian cells. HEK 293T, HeLa, COS-1, and NIH 3T3 cells were transfected with plasmids and RNAs as in Figure 1 and subjected to dual luciferase assays 48 h posttransfection. The ratios of firefly to *Renilla* luciferase activity are normalized to a control transfected with an siRNA directed at the green fluorescent protein (GFP). The average of three independent experiments is shown; error bars indicate standard deviation.

Transcription templates that were predicted to generate siRNAs and shRNAs similar to those prepared by chemical RNA synthesis were prepared by DNA synthesis (Fig. 3A,C). These were tested for efficacy both in S2 cells (data not shown) and in human 293 cells (Fig. 3B,D). Overall, the performance of the T7-synthesized hairpin or siRNAs closely matched the performance of either produced by chemical synthesis, both with respect to the magnitude of inhibition and with respect to the relative efficiency of differing sequences. Because T7 polymerase prefers to initiate at twin guanosine residues, however, it was critical to consider initiation context when designing in vitro transcribed siRNAs (Fig. 3B). In contrast, shRNAs, which are processed by Dicer (see Fig. 1C), tolerate the addition of these bases at the 5′ end of the transcript.

Studies in *Drosophila* embryo extracts indicate that siRNAs possess 5′ phosphorylated termini, consistent with their production by an RNase III family nuclease (Bernstein et al. 2001a; Elbashir et al. 2001b). In vitro, this terminus is critical to the induction of RNAi by synthetic RNA oligonucleotides (Elbashir et al. 2001c; Nykanen et al. 2001). Chemically synthesized siRNAs are nonphosphorylated, and enzymatic addition of a 5′ phosphate group in vitro prior to transfection does not increase the potency of the silencing effect (A.A. Caudy, unpubl.). This suggests either that the requirement for phosphorylated termini is less stringent in mammalian cells or that a kinase efficiently phosphorylates siRNAs in vivo. RNAs synthesized with T7 RNA polymerase, however, possess 5′ triphosphate termini. We therefore explored the possibility of synthesizing siRNAs with T7 polymerase followed by treatment in vitro with pyrophosphatase to modify the termini to resemble those of siRNAs. Surprisingly, monophosphorylated siRNAs (data not shown) were as potent in inducing gene silencing as transcription products bearing triphosphate termini (Fig. 3B). This may suggest either that the requirement for monophosphorylated termini is less stringent in mammalian cells or that siRNAs are modified in vivo to achieve an appropriate terminal structure.

Considered together, our data suggest that both shRNAs and siRNA duplexes can be prepared by synthesis with T7 RNA polymerase in vitro. This significantly reduces the cost of RNAi in mammalian cells and paves the way for application of RNAi on a whole-genome scale.

Transcription of shRNAs in vivo by RNA polymerase III

Although siRNAs are an undeniably effective tool for probing gene function in mammalian cells, their suppressive effects are by definition of limited duration. Delivery of siRNAs can be accomplished by any of a num-

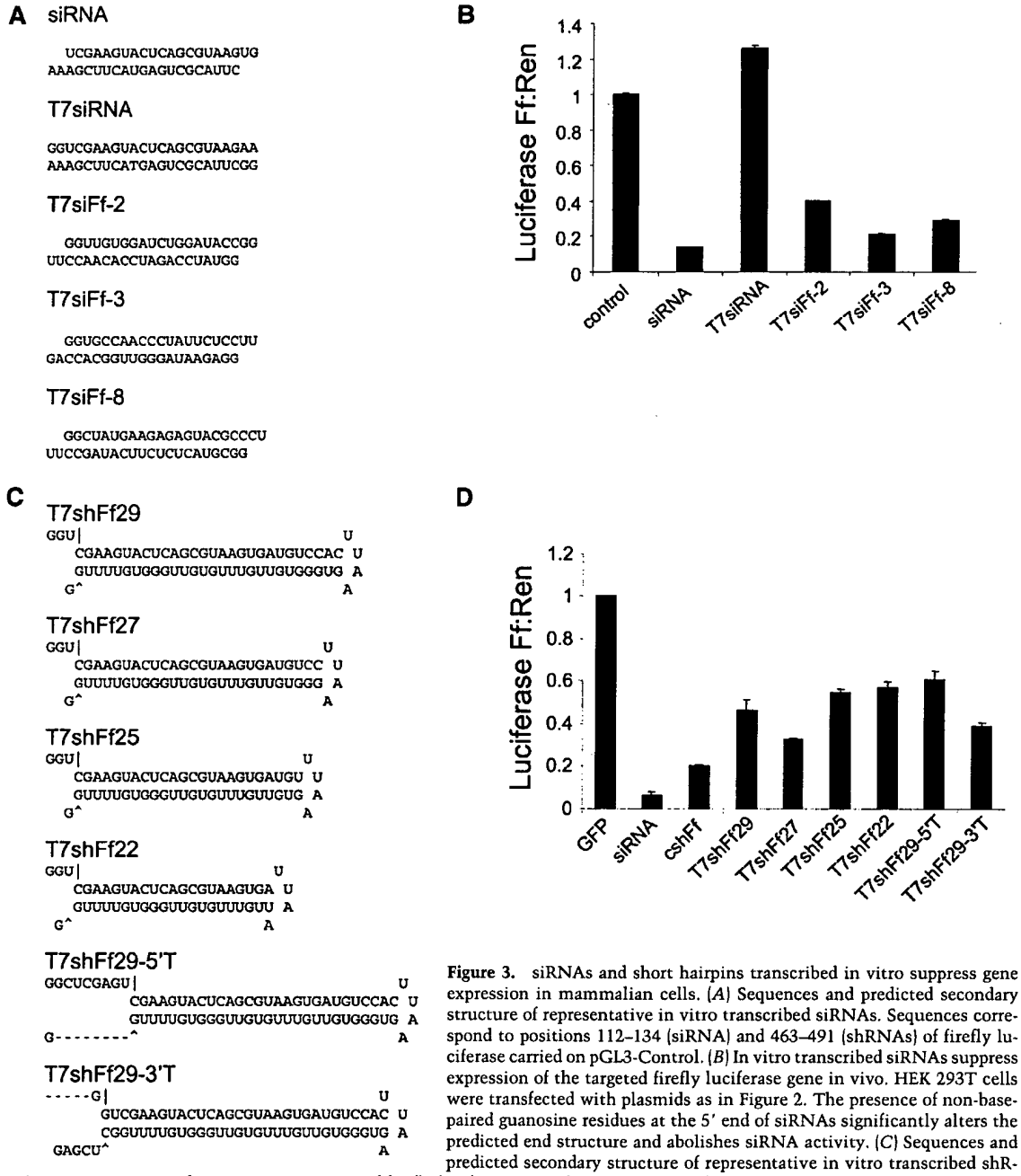


Figure 3. siRNAs and short hairpins transcribed in vitro suppress gene expression in mammalian cells. (*A*) Sequences and predicted secondary structure of representative in vitro transcribed siRNAs. Sequences correspond to positions 112–134 (siRNA) and 463–491 (shRNAs) of firefly luciferase carried on pGL3-Control. (*B*) In vitro transcribed siRNAs suppress expression of the targeted firefly luciferase gene in vivo. HEK 293T cells were transfected with plasmids as in Figure 2. The presence of non-base-paired guanosine residues at the 5′ end of siRNAs significantly alters the predicted end structure and abolishes siRNA activity. (*C*) Sequences and predicted secondary structure of representative in vitro transcribed shRNAs. Sequences correspond to positions 112–141 of firefly luciferase carried on pGL3-Control. (*D*) Short hairpins transcribed in vitro suppress expression of the targeted firefly luciferase gene in vivo. HEK 293T cells were transfected with plasmids as in Figure 2.

ber of transient transfection methodologies, and both the timing of peak suppression and the recovery of protein levels as silencing decays can vary with both the cell type and the target gene (Y. Seger and E. Bernstein, unpubl.). Therefore, one limitation on siRNAs is the development of continuous cell lines in which the expression of a desired target is stably silenced.

Hairpin RNAs, consisting of long duplex structures, have been proved as effective triggers of stable gene silencing in plants, in *C. elegans*, and in *Drosophila* (Ken-

nerdell and Carthew 2000; Smith et al. 2000; Tavernarakis et al. 2000). We have recently shown stable suppression of gene expression in cultured mammalian cells by continuous expression of a long hairpin RNA (Paddison et al. 2002). However, the scope of this approach was limited by the necessity of expressing such hairpins only in cells that lack a detectable PKR response. In principle, shRNAs could bypass such limitations and provide a tool for evoking stable suppression by RNA in mammalian somatic cells.

To test this possibility, we initially cloned sequences encoding a firefly luciferase shRNA into a CMV-based expression plasmid. This was predicted to generate a capped, polyadenylated RNA polymerase II transcript in which the hairpin was extended on both the 5′ and 3′ ends by vector sequences and poly(A). This construct was completely inert in silencing assays in 293T cells (data not shown).

During our studies on chemically and T7-synthesized shRNAs, we noted that the presence of significant single-stranded extensions (either 5′ or 3′ of the duplex) reduced the efficacy of shRNAs (data not shown). We therefore explored the use of alternative promoter strategies in an effort to produce more defined hairpin RNAs. In particular, RNA polymerase III promoters have well-defined initiation and termination sites and naturally produce a variety of small, stable RNA species. Although many Pol III promoters contain essential elements within the transcribed region, limiting their utility for our purposes; class III promoters use exclusively non-transcribed promoter sequences. Of these, the U6 snRNA promoter and the H1 RNA promoter have been well studied (Lobo et al. 1990; Hannon et al. 1991; Chong et al. 2001).

By placing a convenient cloning site immediately behind the U6 snRNA promoter, we have constructed pShh-1, an expression vector in which short hairpins are harnessed for gene silencing. Into this vector either of two shRNA sequences derived from firefly luciferase were cloned from synthetic oligonucleotides. These were cotransfected with firefly and *Renilla* luciferase expression plasmids into 293T cells. One of the two encoded shRNAs provoked effective silencing of firefly luciferase without altering the expression of the internal control (Fig. 4C). The second encoded shRNA also produced detectable, albeit weak, repression. In both cases, silencing was dependent on insertion of the shRNA in the correct orientation with respect to the promoter (Fig. 4C; data not shown). Although the shRNA itself is bilaterally symmetric, insertion in the incorrect orientation would affect Pol III termination and is predicted to produce a hairpin with both 5′ and 3′ single-stranded extensions. Similar results were also obtained in a number of other mammalian cell lines including HeLa, COS-1, NIH 3T3, and IMR90 (Fig. 4; data not shown). pShh1-Ff1 was, however, incapable of effecting suppression of the luciferase reporter in *Drosophila* cells, in which the human U6 promoter is inactive (data not shown).

As a definitive test of whether the plasmid-encoded shRNAs brought about gene silencing via the mammalian RNAi pathway, we assessed the dependence of suppression on an essential component of the RNAi pathway. We transfected pShh1-Ff1 along with an siRNA homologous to human *Dicer*. Figure 5 shows that treatment of cells with *Dicer* siRNAs is able to completely depress the silencing induced by pShh1-Ff1. Addition of an unrelated siRNA had no effect on the magnitude of suppression by pShh1-Ff1 (data not shown). Importantly, *Dicer* siRNAs had no effect on siRNA-induced silencing of firefly luciferase (data not shown). These results are consistent with shRNAs operating via an RNAi pathway similar to those provoked by stRNAs and long dsRNAs. Furthermore, it suggests that siRNA-mediated silencing is less sensitive to depletion of the Dicer enzyme.

The ultimate utility of encoded short hairpins will be in the creation of stable mutants that permit the study of the resulting phenotypes. We therefore tested whether we could create a cellular phenotype through stable suppression. Expression of activated alleles of the *ras* oncogene in primary mouse embryo fibroblasts (MEFs) induces a stable growth arrest that resembles, as a terminal phenotype, replicative senescence (Serrano et al. 1997). Cells cease dividing and assume a typical large, flattened morphology. Senescence can be countered by mutations that inactivate the p53 tumor suppressor pathway (Serrano et al. 1997). As a test of the ability of vector-encoded shRNAs to stably suppress an endogenous cellular gene, we generated a hairpin that was targeted to the mouse *p53* gene. As shown in Figure 6, MEFs transfected with pBabe-RasV12 fail to proliferate and show a senescent morphology when cotransfected with an empty control vector. As noted previously (Serrano et al. 1997), the terminally arrested state is achieved in 100% of drug-selected cells in culture by 8 d posttransfection. However, upon cotransfection of an activated *ras* expression construct with the pShh-p53, cells emerged from drug selection that not only fail to adopt a senescent morphology but also maintain the ability to proliferate for a minimum of several weeks in culture (Fig. 6). These data strongly suggest that shRNA expression constructs can be used for the creation of continuous mammalian cell lines in which selected target genes are stably suppressed.

Discussion

The demonstration that short dsRNA duplexes can induce sequence-specific silencing in mammalian cells has begun to foment a revolution in the manner in which gene function is examined in cultured mammalian cells. These siRNAs (Elbashir et al. 2001a) mimic the products generated by Dicer (Bernstein et al. 2001a) in the initiation step of RNAi and presumably enter the silencing pathway without triggering nonspecific translational suppression via PKR. siRNAs can be used to examine the consequences of reducing the function of virtually any protein-coding gene and have proved effective in provoking relevant phenotypes in numerous somatic cell types from both humans and mice. However, a significant dis-

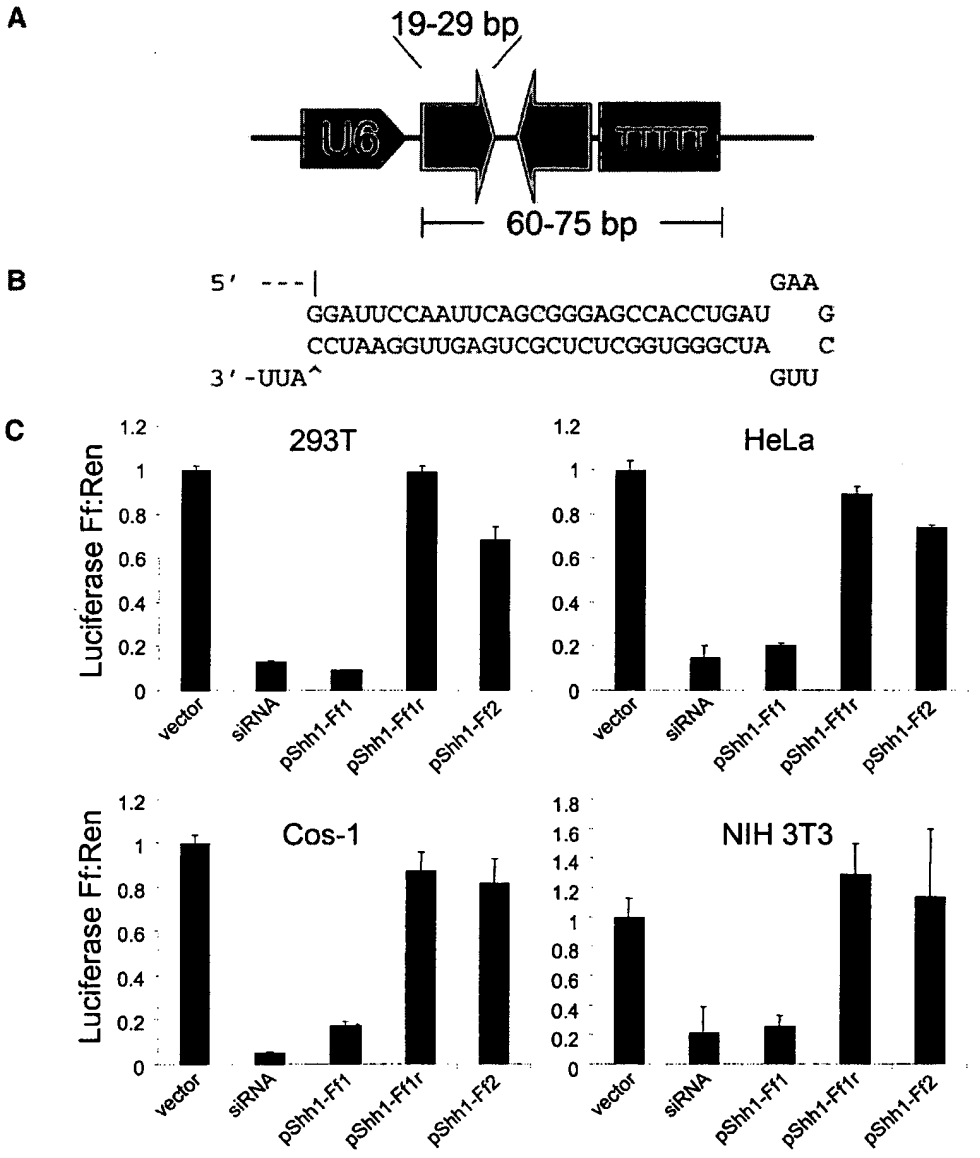


Figure 4. Transcription of functional shRNAs in vivo. (*A*) Schematic of the pShh1 vector. Sequences encoding shRNAs with between 19 and 29 bases of homology to the targeted gene are synthesized as 60–75-bp double-stranded DNA oligonucleotides and ligated into an *Eco*RV site immediately downstream of the U6 promoter. (*B*) Sequence and predicted secondary structure of the Ff1 hairpin. (*C*) An shRNA expressed from the pShh1 vector suppresses luciferase expression in mammalian cells. HEK 293T, HeLa, COS-1, and NIH 3T3 cells were transfected with reporter plasmids as in Figure 1, and pShh1 vector, firefly siRNA, or pShh1 firefly shRNA constructs as indicated. The ratios of firefly to *Renilla* luciferase activity were determined 48 h after transfection and represent the average of three independent experiments; error bars indicate standard deviation.

advantage of siRNAs is that their effects are transient, with phenotypes generated by transfection with such RNAs persisting for ~1 wk. In *C. elegans*, RNAi has proved to be such a powerful tool, in part, because silencing is both systemic and heritable, permitting the consequences of altering gene expression to be examined throughout the development and life of an animal. We have therefore sought to expand the utility of RNAi in mammalian systems by devising methods to induce stable and heritable gene silencing. Previously, we have shown that expression of long (~500-nt) dsRNAs could produce stable silencing in embryonic mammalian cells

Figure 5. Dicer is required for shRNA-mediated gene silencing. HEK 293T cells were transfected with luciferase reporter plasmids as well as pShh1-Ff1 and an siRNA targeting human Dicer either alone or in combination, as indicated. The Dicer siRNA sequence (TCA ACC AGC CAC TGC TGG A) corresponds to coordinates 3137–3155 of the human *Dicer* sequence. The ratios of firefly to *Renilla* luciferase activity were determined 26 h after transfection and represent the average of three independent experiments; error bars indicate standard deviation.

(Paddison et al. 2002); however, the utility of this approach was limited by its restriction to cells that lack endogenous, nonspecific responses to dsRNA, such as PKR.

Recently, a number of laboratories (Grishok et al. 2001; Hutvagner et al. 2001; Ketting et al. 2001; Knight and Bass 2001) have shown that there exist endogenously encoded triggers of RNAi-related pathways, which are transcribed as short hairpin RNAs (stRNAs, or generically miRNAs). Here, we have shown that short hairpin RNAs, modeled conceptually on miRNAs, are potent experimental tools for inducing gene silencing in mammalian somatic cells. These shRNAs can be provided exogenously or can be synthesized in vivo from RNA polymerase III promoters. Not only does this enable the creation of continuous cell lines in which suppression of a target gene is stably maintained by RNAi, but similar strategies may also be useful for the construction of transgenic animals. Thus, short-hairpin-activated gene silencing (SHAGging) provides a complement to the use of siRNAs in the study of gene function in mammalian cells. Finally, the ability to encode a constitutive silencing signal may permit the marriage of shRNA-induced silencing with in vivo and ex vivo gene delivery methods for therapeutic approaches based on stable RNAi in humans.

Materials and methods

Cell culture

HEK 293T, HeLa, COS-1, MEF, and IMR90 cells were cultured in DMEM (GIBCO BRL) supplemented with 10% heat-inactivated fetal bovine serum (FBS) and 1% antibiotic/antimycotic solution (GIBCO BRL). NIH 3T3 cells were cultured in DMEM supplemented with 10% heat-inactivated calf serum and 1% antibiotic/antimycotic solution.

RNA preparation

Both shRNAs and siRNAs were produced in vitro using chemically synthesized DNA oligonucleotide templates (Sigma) and the T7 Megashortscript kit (Ambion). Transcription templates

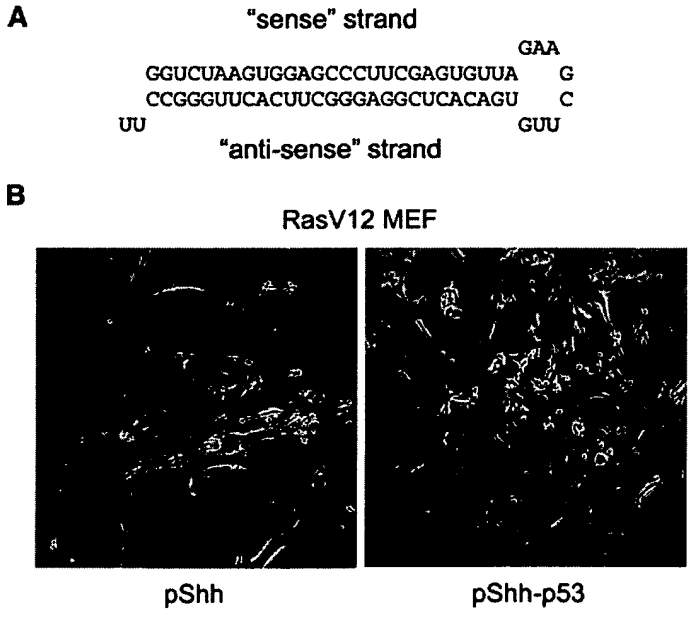


Figure 6. Stable shRNA-mediated gene silencing of an endogenous gene. (*A*) Sequence and predicted secondary structure of the *p53* hairpin. The 5′ shRNA stem contains a 27-nt sequence derived from mouse *p53* (nucleotides 166–192), whereas the 3′ stem harbors the complimentary antisense sequence. (*B*) Senescence bypass in primary mouse embryo fibroblasts (MEFs) expressing an shRNA targeted at *p53*. Wild-type MEFs, passage 5, were transfected with pBabe-RasV12 with control plasmid or with p53hp (5 μg each with FuGENE; Roche). Two days after transfection, cells were trypsinized, counted, and plated at a density of 1×10^5/10-cm plate in media containing 2.0 μg/mL of puromycin. Control cells cease proliferation and show a senescent morphology (*left* panel). Cells expressing the *p53* hairpin continue to grow (*right* panel). Photos were taken 14 d posttransfection.

were designed such that they contained *T7* promoter sequences at the 5′ end. shRNA transcripts subjected to in vitro Dicer processing were synthesized using a Riboprobe kit (Promega). Chemically synthesized RNAs were obtained from Dharmacon, Inc.

Transfection and gene silencing assays

Cells were transfected with indicated amounts of siRNA, shRNA, and plasmid DNA using standard calcium phosphate procedures at 50%–70% confluence in 6-well plates. Dual luciferase assays (Promega) were carried out by cotransfecting cells with plasmids containing firefly luciferase under the control of the SV40 promoter (pGL3-Control, Promega) and *Renilla* luciferase under the control of the SV40 early enhancer/promoter region (pSV40, Promega). Plasmids were cotransfected using a 1:1 ratio of pGL3-Control (250 ng/well) to pRL-SV40. RNAi in S2 cells was performed as previously described (Hammond et al. 2000). For stable silencing, primary MEFs (a gift from S. Lowe, Cold Spring Harbor Laboratory, NY) were cotransfected using Fugene 6 with pBabe-Ha-rasV12 and pShh-p53 (no resistance marker), according to the manufacturer's recommendations. Selection was for the presence of the activated *HarasV12* plasmid, which carries a puromycin-resistance marker. The pShh-p53 plasmid was present in excess, as is standard in a cotransfection experiment. We have now generated a version of the U6 promoter vector (pSHAG-1) that is compatible with the GATEWAY system (Invitrogen), and this can be used to transport the shRNA expression cassette into a variety of recipient vectors that carry *cis*-linked selectable markers. Furthermore, we have validated delivery of shRNAs using retroviral vectors. Updated plasmid information can be obtained at http://www.cshl.org/public/science/hannon.html.

Plasmids expressing hairpin RNAs

The U6 promoter region from −265 to +1 was amplified by PCR, adding 5′ *Kpn*I and 3′ *Eco*RV sites for cloning into pBSSK+. A linker/terminator oligonucleotide set bearing the U6 terminator sequence and linker ends of 5′ *Eco*RV and 3′ *Not*I was cloned into the promoter construct, resulting in a U6 cassette with an *Eco*RV site for insertion of new sequences. This vector has been named pShh1. Blunt-ended, double-stranded DNA oligonucleotides encoding shRNAs with between 19 and 29 bases of homology to the targeted gene were ligated into the *Eco*RV site to produce expression constructs. The oligonucleotide sequence used to construct Ff1 was: TCCAATTCAGCGGGAGCCACC TGATGAAGCTTGATCGGGTGGCTCTCGCTGAGTTGGA ATCCATTTTTTTT. This sequence is preceded by the sequence GGAT, which is supplied by the vector, and contains a tract of more than five Ts as a Pol III terminator.

In vitro Dicer assays

In vitro assays for Dicer activity were performed as described (Bernstein et al. 2001a).

Acknowledgments

We thank members of the Hannon laboratory for critical reading of the manuscript. P.J.P. is an Arnold and Mabel Beckman Anderson Fellow of the Watson School of Biological Sciences, and thanks Richard M. Paddison for academic support. A.A.C. is a George A. and Marjorie H. Anderson Fellow of the Watson School of Biological Sciences and a predoctoral fellow of the Howard Hughes Medical Institute. G.J.H. is a Rita Allen Foundation scholar. This work was supported in part by a grant from the NIH (RO1-GM62534, GJH) and by a grant from Genetica, Inc. (Cambridge, MA).

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J. 2001a. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409:** 363–366.

Bernstein, E., Denli, A.M., and Hannon, G.J. 2001b. The rest is silence. *RNA* **7:** 1509–1521.

Billy, E., Brondani, V., Zhang, H., Muller, U., and Filipowicz, W. 2001. Specific interference with gene expression induced by long, double-stranded RNA in mouse embryonal teratocarcinoma cell lines. *Proc. Natl. Acad. Sci.* **98:** 14428–14433.

Caplen, N.J., Parrish, S., Imani, F., Fire, A., and Morgan, R.A. 2001. Specific inhibition of gene expression by small double-stranded RNAs in invertebrate and vertebrate systems. *Proc. Natl. Acad. Sci.* **98:** 9742–9747.

Chong, S.S., Hu, P., and Hernandez, N. 2001. Reconstitution of transcription from the human U6 small nuclear RNA promoter with eight recombinant polypeptides and a partially purified RNA polymerase III complex. *J. Biol. Chem.* **276:** 20727–20734.

Clarke, P.A. and Mathews, M.B. 1995. Interactions between the double-stranded RNA binding motif and RNA: Definition of the binding site for the interferon-induced protein kinase DAI (PKR) on adenovirus VA RNA. *RNA* **1:** 7–20.

Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. 2001a. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411:** 494–498.

Elbashir, S.M., Lendeckel, W., and Tuschl, T. 2001b. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes & Dev.* **15:** 188–200.

Elbashir, S.M., Martinez, J., Patkaniowska, A., Lendeckel, W., and Tuschl, T. 2001c. Functional anatomy of siRNAs for mediating efficient RNAi in *Drosophila melanogaster* embryo lysate. *EMBO J.* **20:** 6877–6888.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391:** 806–811.

Gil, J. and Esteban, M. 2000. Induction of apoptosis by the dsRNA-dependent protein kinase (PKR): Mechanism of action. *Apoptosis* **5:** 107–114.

Grishok, A., Pasquinelli, A.E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D.L., Fire, A., Ruvkun, G., and Mello, C.C. 2001. Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* **106:** 23–34.

Guo, S. and Kemphues, K.J. 1995. par-1, a gene required for establishing polarity in *C. elegans* embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed. *Cell* **81:** 611–620.

Ha, I., Wightman, B., and Ruvkun, G. 1996. A bulged lin-4/lin-14 RNA duplex is sufficient for *Caenorhabditis elegans* lin-14 temporal gradient formation. *Genes & Dev.* **10:** 3041–3050.

Hammond, S.M., Bernstein, E., Beach, D., and Hannon, G.J.

2000. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* **404:** 293–296.

Hammond, S.M., Boettcher, S., Caudy, A.A., Kobayashi, R., and Hannon, G.J. 2001a. Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science* **293:** 1146–1150.

Hammond, S.M., Caudy, A.A., and Hannon, G.J. 2001b. Post-transcriptional gene silencing by double-stranded RNA. *Nat. Rev. Genet.* **2:** 110–119.

Hannon, G.J., Chubb, A., Maroney, P.A., Hannon, G., Altman, S., and Nilsen, T.W. 1991. Multiple cis-acting elements are required for RNA polymerase III transcription of the gene encoding H1 RNA, the RNA component of human RNase P. *J. Biol. Chem.* **266:** 22796–22799.

Hutvagner, G., McLachlan, J., Pasquinelli, A.E., Balint, E., Tuschl, T., and Zamore, P.D. 2001. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* **293:** 834–838.

Kennerdell, J.R. and Carthew, R.W. 2000. Heritable gene silencing in *Drosophila* using double-stranded RNA. *Nat. Biotechnol.* **18:** 896–898.

Ketting, R.F., Fischer, S.E., Bernstein, E., Sijen, T., Hannon, G.J., and Plasterk, R.H. 2001. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes & Dev.* **15:** 2654–2659.

Knight, S.W. and Bass, B.L. 2001. A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*. *Science* **293:** 2269–2271.

Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. 2001. Identification of novel genes coding for small expressed RNAs. *Science* **294:** 853–858.

Lau, N.C., Lim, L.P., Weinstein, E.G., and Bartel, D.P. 2001. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* **294:** 858–862.

Lee, R.C. and Ambros, V. 2001. An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* **294:** 862–864.

Lobo, S.M., Ifill, S., and Hernandez, N. 1990. cis-Acting elements required for RNA polymerase II and III transcription in the human U2 and U6 snRNA promoters. *Nucleic Acids Res.* **18:** 2891–2899.

Nykanen, A., Haley, B., and Zamore, P.D. 2001. ATP requirements and small interfering RNA structure in the RNA interference pathway. *Cell* **107:** 309–321.

Paddison, P., Caudy, A.A., and Hannon, G.J. 2002. Stable suppression of gene expression in mammalian cells by RNAi. *Proc. Natl. Acad. Sci.* **99:** 1443–1448.

Reinhart, B.J., Slack, F.J., Basson, M., Pasquinelli, A.E., Bettinger, J.C., Rougvie, A.E., Horvitz, H.R., and Ruvkun, G. 2000. The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* **403:** 901–906.

Serrano, M., Lin, A.W., McCurrach, M.E., Beach, D., and Lowe, S.W. 1997. Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. *Cell* **88:** 593–602.

Shi, H., Djikeng, A., Mark, T., Wirtz, E., Tschudi, C., and Ullu, E. 2000. Genetic interference in *Trypanosoma brucei* by heritable and inducible double-stranded RNA. *RNA* **6:** 1069–1076.

Slack, F.J., Basson, M., Liu, Z., Ambros, V., Horvitz, H.R., and Ruvkun, G. 2000. The lin-41 RBCC gene acts in the *C. elegans* heterochronic pathway between the let-7 regulatory RNA and the LIN-29 transcription factor. *Mol. Cell* **5:** 659–669.

Smith, N.A., Singh, S.P., Wang, M.B., Stoutjesdijk, P.A., Green, A.G., and Waterhouse, P.M. 2000. Total silencing by intron-spliced hairpin RNAs. *Nature* **407:** 319–320.

Svoboda, P., Stein, P., Hayashi, H., and Schultz, R.M. 2000. Selective reduction of dormant maternal mRNAs in mouse oocytes by RNA interference. *Development* **127:** 4147–4156.

Tavernarakis, N., Wang, S.L., Dorovkov, M., Ryazanov, A., and Driscoll, M. 2000. Heritable and inducible genetic interference by double-stranded RNA encoded by transgenes. *Nat. Genet.* **24:** 180–183.

Tuschl, T., Zamore, P.D., Lehmann, R., Bartel, D.P., and Sharp, P.A. 1999. Targeted mRNA degradation by double-stranded RNA in vitro. *Genes & Dev.* **13:** 3191–3197.

Wianny, F. and Zernicka-Goetz, M. 2000. Specific interference with gene function by double-stranded RNA in early mouse development. *Nat. Cell Biol.* **2:** 70–75.

Wightman, B., Ha, I., and Ruvkun, G. 1993. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in *C. elegans*. *Cell* **75:** 855–862.

Williams, B.R. 1997. Role of the double-stranded RNA-activated protein kinase (PKR) in cell regulation. *Biochem. Soc. Trans.* **25:** 509–513.

Yang, S., Tutton, S., Pierce, E., and Yoon, K. 2001. Specific double-stranded RNA interference in undifferentiated mouse embryonic stem cells. *Mol. Cell. Biol.* **21:** 7807–7816.

Stable suppression of gene expression by RNAi in mammalian cells

Patrick J. Paddison*, Amy A. Caudy*, and Gregory J. Hannon†

Watson School of Biological Sciences, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724

Communicated by Bruce W. Stillman, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, December 6, 2001 (received for review October 2, 2001)

In a diverse group of organisms including plants, *Caenorhabditis elegans*, *Drosophila*, and trypanosomes, double-stranded RNA (dsRNA) is a potent trigger of gene silencing. In several model systems, this natural response has been developed into a powerful tool for the investigation of gene function. Use of RNA interference (RNAi) as a genetic tool has recently been extended to mammalian cells, being inducible by treatment with small, ≈22-nt RNAs that mimic those produced in the first step of the silencing process. Here, we show that some cultured murine cells specifically silence gene expression upon treatment with long dsRNAs (≈500 nt). This response shows hallmarks of conventional RNAi including silencing at the posttranscriptional level and the endogenous production of ≈22-nt small RNAs. Furthermore, enforced expression of long, hairpin dsRNAs induced stable gene silencing. The ability to create stable "knock-down" cell lines expands the utility of RNAi in mammalian cells by enabling examination of phenotypes that develop over long time periods and lays the groundwork for by using RNAi in phenotype-based, forward genetic selections.

The use of genetically tractable model systems has been the key to our present understanding of gene structure and function, cell and organismal biology, and, ultimately, the molecular aspects of human disease. The ability to stably knock out or knock down gene expression and, thus, function, in particular, has been paramount to the use of such models for illuminating biological function. For example, the use of conditional lethals in bacteriophage T4 allowed functional analysis of phage morphogenesis modules (1), whereas the same technique applied to yeast permitted the discovery of functional hierarchies among genes regulating cell cycle progression (2, 3). In both scenarios, cells acquire stable phenotypes through heritable genetic alterations.

Although such basic genetic approaches are virtually effortless in many model organisms, cultured mammalian cells have proven somewhat intractable, in this regard. This is largely because cultured mammalian cells are diploid and favor nonhomologous over homologous recombination. Current approaches to create stable phenotypes in mammalian cells have been often met with limited success. Dominant-negative and antisense strategies have proven inconsistent and unpredictable, thus lacking experimental rigor equivalent to a point mutation in yeast. However, one approach now used extensively in other diploid organisms has the potential to foment a revolution in mammalian somatic cell genetics. This approach is dubbed double-stranded RNA (dsRNA)–dependent posttranscriptional gene silencing, or RNA interference (RNAi).

It has become clear that dsRNA-induced silencing phenomena are present in evolutionarily diverse organisms including plants, fungi, and metazoans (reviewed in ref. 4). A combination of genetic and biochemical studies suggests that many of these phenomena share a common mechanism. The prevailing model begins with the conversion of the dsRNA silencing "trigger" into small RNAs (guide RNAs or siRNAs, ref. 5) that range in size from ≈21 to 25 nts, depending on the species of origin (6–8). These RNAs become incorporated into a multicomponent nuclease complex, which uses the sequence of the guide/siRNAs to identify and destroy homologous mRNAs (7, 8).

In several systems, dsRNA-induced silencing has been harnessed as a powerful tool for the analysis of gene function. Particularly in *Caenorhabditis elegans*, RNAi has emerged as the standard protocol for quickly assessing the consequences of inhibiting gene function. In fact, programs are underway to create RNAi libraries that can be used to suppress, individually, each of the ≈19,000 genes in the worm genome (9, 10). In *Drosophila*, the first evidence of dsRNA-induced silencing came from the study of embryos (11), and subsequently, RNAi has proven an effective tool in cultured cells and in adult insects (7, 12, 13).

Despite its utility in diverse systems, harnessing RNA to study gene function in mammals seemed potentially problematic. Indeed, mammals have evolved robust systems for responding to dsRNAs, specifically as an antiviral defense (reviewed in refs. 14 and 15). In somatic cells, dsRNA activates a variety of responses. Predominant among these is PKR, a kinase that is activated by dimerization in the presence of dsRNA (16). PKR, in turn, phosphorylates EIF2α, causing a nonspecific translational shutdown (reviewed in ref. 14). dsRNA also activates 2′-5′ oligoadenylate polymerase, the product of which is an essential cofactor for a nonspecific ribonuclease, RNase L (reviewed in ref. 17).

Recently, Tuschl and colleagues (5) have demonstrated that RNAi can be provoked in numerous mammalian cell lines through the introduction of siRNAs. These siRNAs avoid provoking the PKR response by virtue of their small size and are presumed to be incorporated into the RNAi pathway by mimicking the products of the Dicer enzyme, which catalyzes the initiation step of RNAi (18). The ability to apply RNAi in mammals will undoubtedly spark a firestorm of effort to assess the consequences of suppressing the expression of genes in cultured mammalian cells.

The power of RNAi as a genetic tool would be greatly enhanced by the ability to engineer stable silencing of gene expression. Whereas the production of small RNAs via *in vivo* expression is problematic, stable silencing has been induced in model organisms by directed expression of long dsRNAs (13, 19, 20). We therefore undertook an effort to identify mammalian cells in which long dsRNAs could be used as RNAi triggers in the hope that these same cell lines would provide a platform upon which to develop stable silencing strategies.

Materials and Methods

Cell Culture. P19 mouse embryonic carcinoma cells (American Type Culture Collection, CRL-1825) were cultured in α-MEM (GIBCO/BRL) supplemented with 10% heat-inactivated FBS and 1% antibiotic/antimycotic solution (GIBCO/BRL). Mouse embryo stem cells (J1, provided by S. Kim, Cold Spring Harbor Laboratory) were cultured in DMEM containing ESgro (Chemicon) according to the manufacturer's instructions. C2C12 murine myoblast cells (gift of N. Tonks, Cold Spring Harbor

Abbreviations: dsRNA, double-stranded RNA; RNAi, RNA interference; siRNA, small interfering RNA; EGFP, enhanced green fluorescent protein.

*P.J.P. and A.A.C. contributed equally to this work.

†To whom reprint requests should be addressed. E-mail: hannon@cshl.org.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

GENETICS

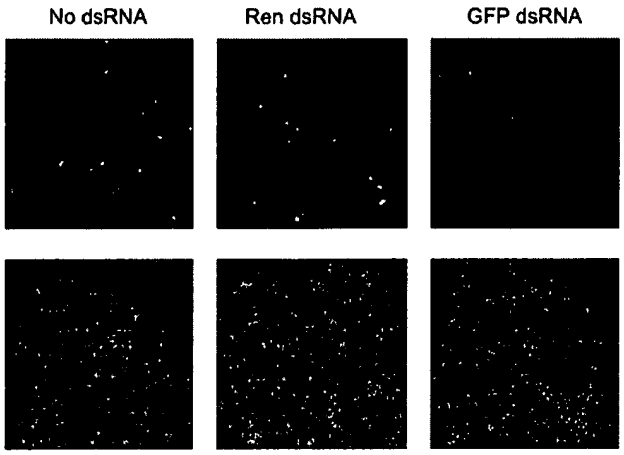


Fig. 1. RNAi in P19 embryonal carcinoma cells. Ten-centimeter plates of P19 cells were transfected by using 5 μg of GFP plasmid and 40 μg of the indicated dsRNA (or no RNA). Cells were photographed by fluorescent and phase-contrast microscopy at 72 h after transfection; silencing was also clearly evident at 48 h posttransfection.

Laboratory) were cultured in DMEM (GIBCO/BRL) supplemented with 10% heat-inactivated FBS and 1% antibiotic/antimycotic solution (GIBCO/BRL).

RNA Preparation. For the production of dsRNA, transcription templates were generated by PCR; they contained T7 promoter sequences on each end of the template (see ref. 7). dsRNAs were prepared by using the RiboMax kit (Ambion, Austin, TX). Firefly and *Renilla* luciferase mRNA transcripts were synthesized by using the Riboprobe kit (Promega) and were gel purified before use.

Transfection and Gene Silencing Assays. Cells were transfected with indicated amounts of dsRNA and plasmid DNA by using FuGENE6 (Roche Biochemicals) according to the manufacturer's instructions. Cells were transfected at 50–70% confluence in 12-well plates containing either 1 or 2 ml of medium per well. Dual luciferase assays (Promega) were carried out by cotransfecting cells with plasmids contain firefly luciferase under the control of SV40 promoter (pGL3-Control, Promega) and *Renilla* luciferase under the control of the SV40 early enhancer/promoter region (pSV40, Promega). These plasmids were cotransfected by using a 1:1 or 10:1 ratio of pGL3-control (250 ng/well) to pRL-SV40. Both ratios yielded similar results. For some experiments, cells were transfected with vectors that direct expression of enhanced green fluorescent protein (EGFP)-US9 fusion protein (21) or red fluorescent protein (pDsRed N1, CLONTECH). RNAi in S2 cells was performed as described (7).

Plasmids expressing hairpin RNAs (RNAs with a self-complimentary stem loop) were constructed by cloning the first 500 bp of the EGFP coding region (CLONTECH) into the FLIP cassette of pRIP-FLIP (E. Bernstein and G.J.H., unpublished data) as a direct repeat. The FLIP cassette contains two directional cloning sites, the second of which sports flanking LoxP sites (see Fig. 6*A*). The Zeocin gene (Stratagene), present between the cloning sites, maintains selection and, thus, stability of the FLIP cassette. The FLIP cassette containing EGFP direct repeats was subcloned into pcDNA3 (Invitrogen). To create an inverted repeat for hairpin production, EGFP direct repeat clones were exposed to Cre recombinase (Stratagene) *in vitro* and, afterward, transformed into DL759 *Escherichia coli* (22). These bacteria permit the replication of DNA containing cruciform structures, which tend to form from inverted repeats.

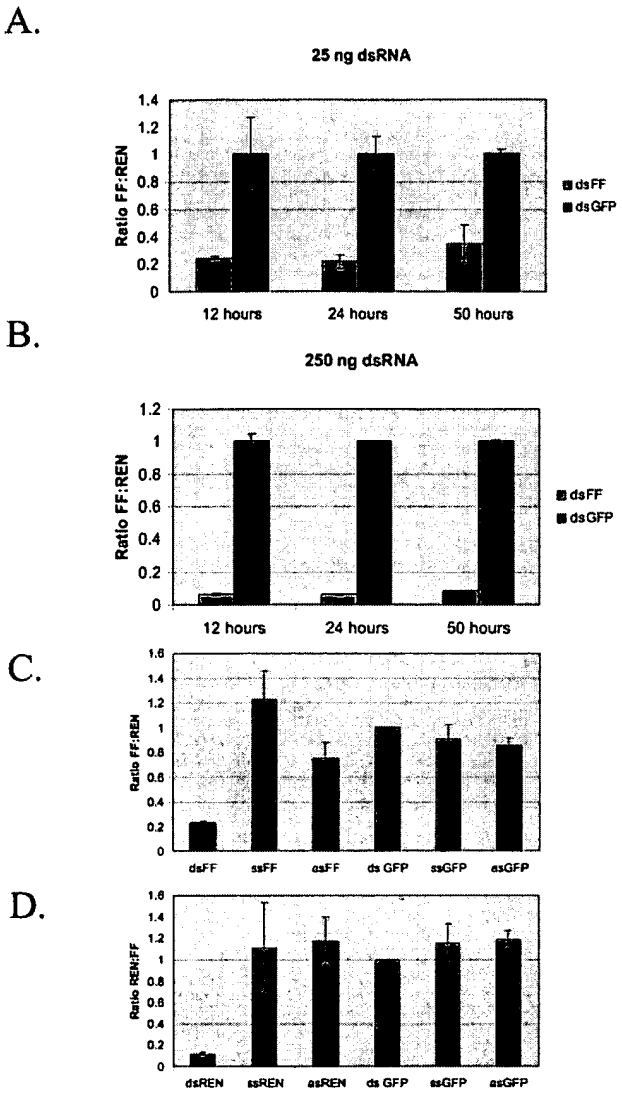


Fig. 2. RNAi of firefly and *Renilla* luciferase in P19 cells. (*A*) P19 cells were transfected with plasmids that direct the expression of firefly and *Renilla* luciferase and dsRNA 500 mers (25 or 250 ng, as indicated), that were either homologous to the firefly luciferase mRNA (dsFF) or nonhomologous (dsGFP). Luciferase activities were assayed at various times after transfection, as indicated. Ratios of firefly to *Renilla* activity are normalized to dsGFP controls. (*B* and *C*) P19 cells in 12-well culture dishes (2 ml of media) were transfected with 0.25 μg of a 9:1 mix of pGL3-Control and pRL-SV40 as well as 2 μg of the indicated RNA. Extracts were prepared 9 h after transfection. (*B*) Ratio of firefly to *Renilla* luciferase is shown. (*C*) Ratio of *Renilla* to firefly luciferase is shown. Values are normalized to dsGFP. The average of three independent experiments is shown; error bars indicate standard deviation.

DL759 transformants were screened for plasmids containing inverted repeats (≈50%).

Silencing of Dicer was accomplished by using a dsRNA comprising exon 25 of the mouse Dicer gene and corresponding to nucleotides 5284–5552 of the human Dicer cDNA.

***In Vitro* Translation and *in Vitro* Dicer Assays.** Logarithmically growing cells were harvested in PBS containing 5 mM EGTA washed twice in PBS and once in hypotonic buffer (10 mM Hepes, pH 7.3/6 mM β-mercaptoethanol). Cells were suspended in 0.7 packed-cell

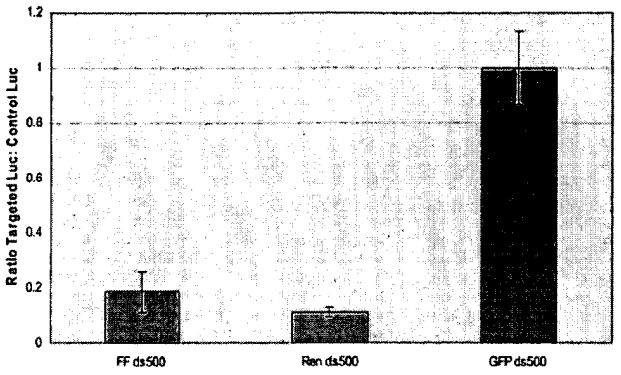


Fig. 3. Specific silencing of luciferase expression by dsRNA in murine embryonic stem cells. Mouse embryonic stem cells in 12-well culture dishes (1 ml of media) were transfected with 1.5 μg of dsRNA along with 0.25 μg of a 10:1 mixture of the reporter plasmids pGL3-Control and pRL-SV40. Extracts were prepared and assayed 20 h after transfection. The ratio of firefly to *Renilla* luciferase expression is shown for FF ds500; the ratio of *Renilla* to firefly is shown for Ren ds500. Both are normalized to ratios from the dsGFP transfection. The average of three independent experiments is shown; error bars indicate standard deviation.

volumes of hypotonic buffer containing *Complete* protease inhibitors (Roche Molecular Biochemicals) and 0.5 units/ml of RNasin (Promega). Cells were disrupted in a Dounce homogenizer with a type B pestle, and lysates were centrifuged at 30,000 × *g* for 20 min. Supernatants were used in an *in vitro* translation assay containing capped m7G(5')pppG firefly and *Renilla* luciferase mRNA or in *in vitro* Dicer assays containing ^{32}P-labeled dsRNA. For *in vitro* translation assays, 5 μl of extract were mixed with 100 ng of firefly and *Renilla* mRNA along with 1 μg of dsRNA (or buffer)/10 mM DTT/0.5 mM spermidine/200 mM Hepes, 3.3 mM MgOAc/800 mM KOAc/1 mM ATP/1 mM GTP/4 units of Rnasin/215 μg of creatine phosphate/1 μg of creatine phosphate kinase/1 mM amino acids (Promega). Reactions were carried out for 1 h at 30°C and quenched by adding 1× passive lysis buffer (Promega). Extracts were then assayed for luciferase activity. *In vitro* assays for Dicer activity were performed as described (18).

Construction of Stable Silencing Lines. Ten-centimeter plates of P19 cells were transfected with 5 μg of GFP hairpin expression plasmid and selected for stable integrants by using G-418 (300 ng/ml) for 14 days. Clones were selected and screened for silencing of GFP.

Results

RNAi in Pluripotent Murine P19 Cells. It has long been clear that the nonspecific responses to dsRNA are attenuated during early development. In fact, injection of dsRNA into early-stage mouse embryos can induce sequence-specific silencing of both exogenous and endogenous genes (23, 24). Consistent with the possibility that RNAi might extend to mammals, homologs of the proteins that participate in this response can be easily identified in the mouse and human genomes (reviewed in ref. 4).

We sought to determine whether long dsRNA triggers could induce sequence-specific silencing in cultured murine cells, both to develop this approach as a tool for probing gene function and to allow mechanistic studies of dsRNA-induced silencing to be propagated to mammalian systems. We, therefore, attempted to extend previous studies in mouse embryos (23, 24) by searching for RNAi-like mechanisms in pluripotent, embryonic cell types.

We surveyed a number of cell lines of embryonic origin for the degree to which generalized suppression of gene expression occurred upon introduction of dsRNA. As an assay, we tested the

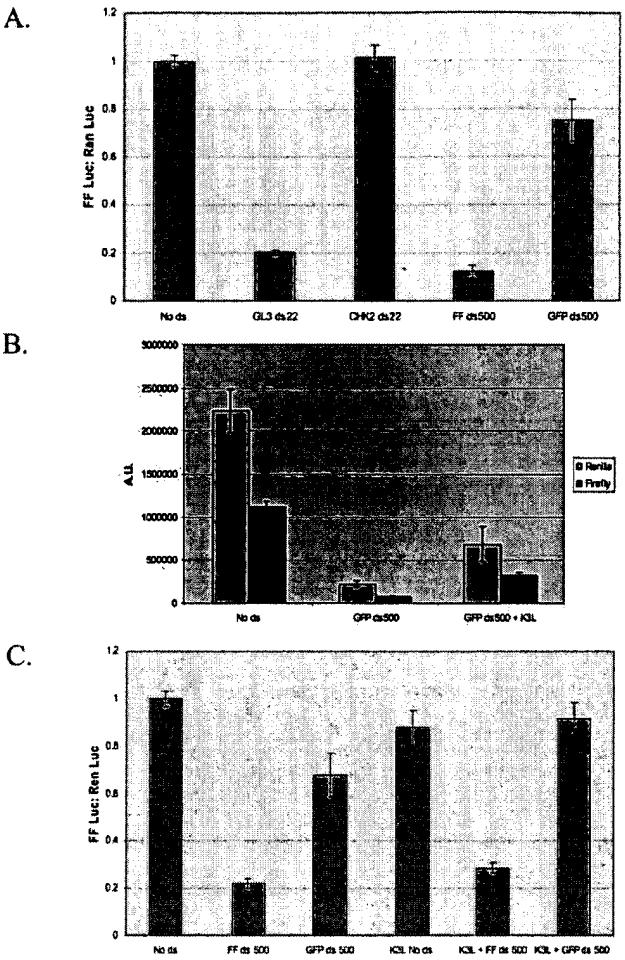


Fig. 4. RNAi in C2C12 murine myoblast cells. (*A*) Mouse C2C12 cells in 12-well culture dishes (1 ml of media) were transfected with 1 μg of the indicated dsRNA along with 0.250 μg of the reporter plasmids pGL3-Control and pRL-SV40. Extracts were prepared and assayed 24 h after transfection. The ratio of firefly to *Renilla* luciferase expression is shown; values are normalized to ratios from the no dsRNA control. The average of three independent experiments is shown; error bars indicate standard deviation. (*B*) C2C12 cells cotransfected with 1 μg of either plasmid alone or a plasmid containing a hyperactive mutant of vaccinia virus K3L (26). The absolute counts of *Renilla* and firefly luciferase activity are shown. (C) The ratios of firefly/*Renilla* activity from *B*, normalized to no dsRNA controls.

effects of dsRNA on the expression of GFP as measured *in situ* by counting fluorescent cells. As expected, in both human embryonic kidney cells (293) and mouse embryo fibroblasts, GFP expression was virtually eliminated irrespective of the sequence of the cotransfected dsRNA (not shown). In some pluripotent teratocarcinoma and teratoma cell lines (e.g., N-Tera1, F9), the PKR response was attenuated but still evident (not shown); however, in contrast, transfection of nonhomologous dsRNAs had no effect on the expression of reporter genes (e.g., GFP, luciferase) either in mouse embryonic stem cells (not shown) or in p19 embryonal carcinoma cells (Fig. 1).

Transfection of P19 embryonal carcinoma cells with GFP in the presence of cognate dsRNA corresponding to the first ≈500 nts of the GFP coding sequence had a strikingly different effect. GFP expression was eliminated in the vast majority of cotransfected cells (Fig. 1), suggesting that these cultured murine cells might respond

GENETICS

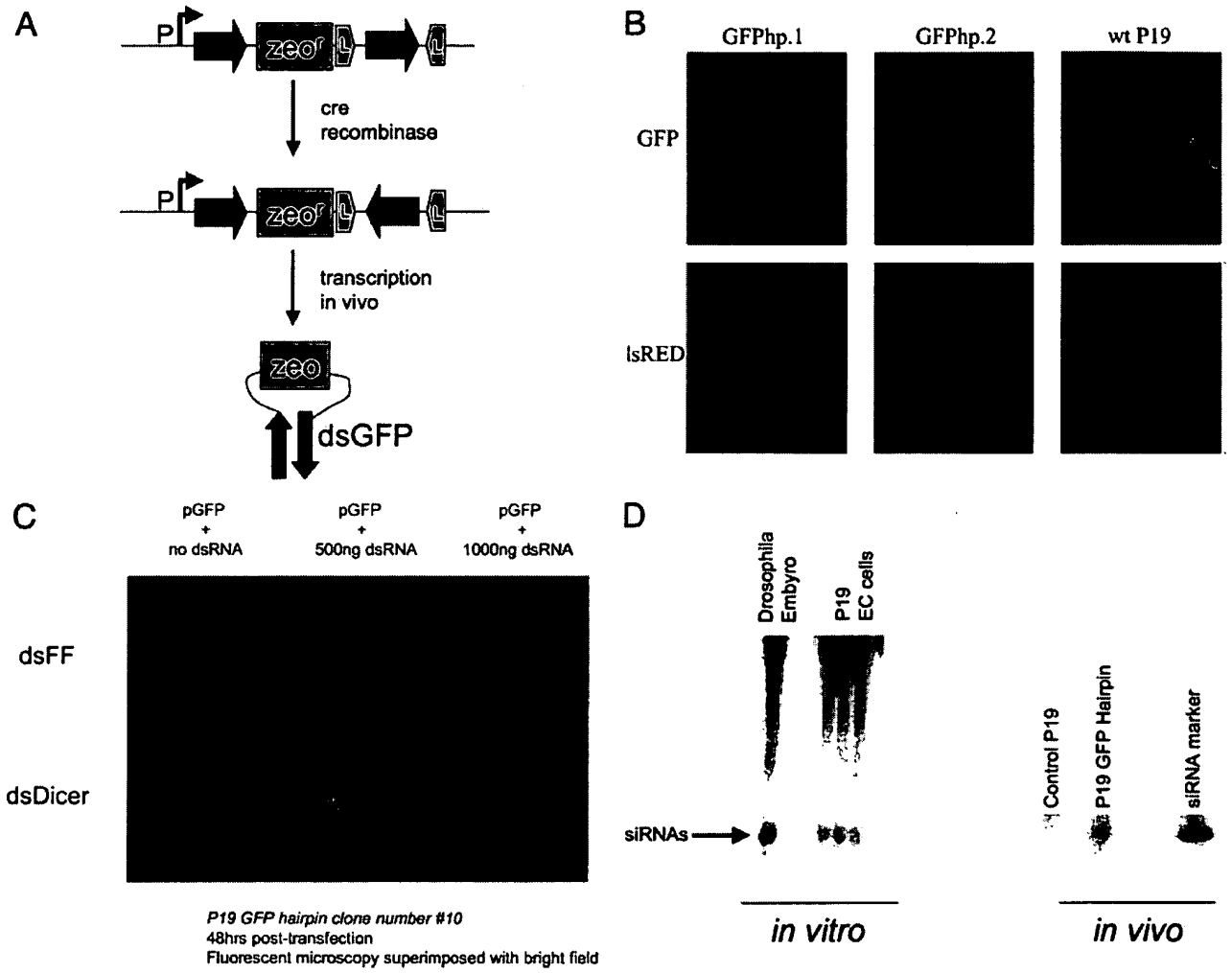


Fig. 5. Expression of a hairpin RNA produces P19 EC cell lines that stably silence GFP. (*A*) A cartoon of the FLIP cassette used to construct the GFP hairpin. GFP represents the first 500 coding base pairs of EGFP. Zeo, zeocin resistance gene; L, Lox; P, the cytomegalovirus promoter in the expression plasmid pcDNA3. Homologous GFP fragments are first cloned as direct repeats into the FLIP cassette. To create inverted repeats for hairpin production, the second repeat is flipped by using Cre recombinase. When transcribed, the inverted repeat forms a GFP dsRNA with a hairpin loop. (*B*) P19 cell lines stably expressing the GFP hairpin plasmid, GFPhp.1 (clone 10) and GFPhp.2 (clone 12), along with wt P19 were transfected with 0.25 μg each of GFP and RFP reporter genes. Fluorescence micrographs were taken by using filters appropriate for GFP and RFP. Magnification is 200×. (*C*) P19 GFPhp.1 cells were transfected with pEGFP and 0, 0.5, or 1 μg of Dicer or firefly dsRNA. Fluorescence micrographs were taken at 48 h posttransfection and are superimposed with bright field images to reveal non-GFP expressing cells. Magnification is 100×. (*D*) *In vitro* and *in vivo* processing of dsRNA in P19 cells. *In vitro* Dicer assays were performed on S2 cells and three independently prepared P19 extracts by using ^{32}P-labeled dsRNA (30°C for 30 min). A Northern blot of RNA extracted from control and GFPhp.1 P19 cells shows the production of ≈22mer RNA species in hairpin-expressing cells but not in control cells. Blots were probed with a ^{32}P-labeled "sense" GFP transcript.

to dsRNA in a manner similar to that which we had previously demonstrated in cultured, *Drosophila* S2 cells (7).

To quantify the extent to which dsRNA could induce sequence-specific gene silencing, we used a dual luciferase reporter assay similar to that which had first been used to demonstrate RNAi in *Drosophila* embryo extracts (25). P19 EC cells were transfected with a mixture of two plasmids that individually direct the expression of firefly luciferase and *Renilla* luciferase. These were cotransfected with no dsRNA, with dsRNA that corresponds to the first ≈500 nts of the firefly luciferase, or with dsRNA corresponding to the first ≈500 nts of GFP as a control. Cotransfection with GFP dsRNA gave luciferase activities that were similar to the no-dsRNA control, both in the firefly/*Renilla* activity ratio and in the absolute values of both activities. In contrast, in cells that received the firefly luciferase dsRNA, the ratio of firefly to *Renilla* luciferase activity was reduced by up to 30-fold (250 ng, Fig. 2*B*). For comparison, we carried out an identical set of experiments in *Drosophila* S2 cells. Although qualitatively similar results were obtained, the silencing response was more potent. At equivalent levels of dsRNA, S2 cells suppressed firefly luciferase activity to virtually background levels (not shown).

The complementary experiment, in which dsRNA was homologous to *Renilla* luciferase, was also performed. Again, in this case, suppression of the expression of the *Renilla* enzyme was ≈10-fold (Fig. 2*D*). Thus, the dsRNA response in P19 cells was flexible, and the silencing machinery was able to adapt to dsRNAs directed against any of the reporters that were tested.

We took two approaches to test whether this response was specific for dsRNA. Pretreatment of the trigger with purified RNase III, a dsRNA-specific ribonuclease, before transfection

greatly reduced its ability to provoke silencing (not shown). Furthermore, transfection of cells with single-stranded antisense RNAs directed against either firefly or *Renilla* luciferase had little or no effect on expression of the reporters (Fig. 2 *C* and *D*). Considered together, these results provided a strong indication that double-stranded RNAs provoke a potent and specific silencing response in P19 embryonal carcinoma cells.

Efficient silencing could be provoked with relatively low concentrations of dsRNA (25 ng/ml culture media; see Fig. 2*A*). The response was concentration-dependent, with maximal suppression of ≈20-fold being achieved at a dose of 1.5 μg/ml culture media.

Silencing was established rapidly and was evident by 9 h post-transfection (the earliest time point examined). Furthermore, the response persisted without significant changes in the degree of suppression for up to 72 h following a single dose of dsRNA.

RNAi in Embryonic Stem Cells. To assess whether the presence of a sequence-specific response to dsRNA was a peculiarity of P19 cells or whether it also extended to normal murine embryonic cells, we performed similar silencing assays in mouse embryonic stem cells. Cotransfection of embryonic stem cells with noncognate dsRNAs (e.g., GFP), again, had no dramatic effect on either the absolute values or the ratios of *Renilla* and firefly luciferase activity (Fig. 3). However, transfection with either firefly or *Renilla* luciferase dsRNA dramatically and specifically reduced the activity of the targeted enzyme (Fig. 3).

This result suggests that RNAi can operate in multiple murine cell types of embryonic origin, including normal embryonic stem cells. The ability to provoke silencing in a cell type that is normally used for the generation of genetic, mosaic animals suggests the possibility of eventually testing the biological effects of silencing both in culture and in reconstituted animal models.

RNAi in Murine Somatic Cells. RNAi effector pathways are likely to be present in mammalian somatic cells, based on the ability of siRNAs to induce transient silencing (5). Furthermore, we have shown that RNAi initiator and effector pathways clearly exist in embryonic cells that can enforce silencing in response to long dsRNA triggers. We therefore sought to test whether the RNAi machinery might exist intact in some somatic cell lines.

Transfection of HeLa cells with luciferase reporters in combination with long dsRNA triggers caused a nearly complete suppression of activity, irrespective of the RNA sequence. In a murine myoblast cell line, C2C12, we noted a mixture of two responses. dsRNAs homologous to firefly luciferase provoked a sequence-specific effect, producing a degree of suppression that was slightly more potent than was observed upon transfection with cognate ≈21-nt siRNA (ref. 5; Fig. 4*A*). However, with long dsRNA triggers, the specific effect was superimposed upon a generalized suppression of reporter gene expression that was presumably because of PKR activation (Fig. 4*B*).

Numerous mammalian viruses have evolved the ability to block PKR as an aid to efficient infection. For example, adenoviruses express VA RNAs, which mimic dsRNA with respect to binding but not to activation of PKR (16). Vaccinia virus uses two strategies to evade PKR. First is expression of E3L, which binds and masks dsRNAs (26). The second is expression of K3L, which binds and inhibits PKR via its ability to mimic the natural substrate of this enzyme, eIF2α (26).

Transfection of C2C12 cells with a vector that directs K3L expression attenuates the generalized repression of reporter genes in response to dsRNA. However, this protein had no effect on the magnitude of specific inhibition by RNAi (Fig. 4*C*).

These results raise the possibility that, at least in some cell lines and/or cell types, blocking nonspecific responses to dsRNA will enable the use of long dsRNAs for the study of gene function. This might be accomplished through the use of viral inhibitors, as described here, or through the use of cells isolated from animals that are genetically modified to lack undesirable responses.

Stable Suppression of Gene Expression Using RNAi. To date, dsRNAs have been used to induce sequence-specific gene silencing in either cultured mammalian cells or in embryos only in a transient fashion. However, the most powerful applications of genetic manipulation are realized only with the creation of stable mutants. The ability to induce silencing by using long dsRNAs offers the opportunity to translate into mammalian cells work from model systems such as *Drosophila*, plants, and *C. elegans* wherein stable silencing has been achieved by enforced expression of hairpin RNAs (13, 19, 20).

P19 EC cells were transfected with a control vector or with an expression vector that directs expression of a ≈500-nt GFP hairpin RNA from an RNA polymerase II promoter (cytomegalovirus). Colonies arising from cells that had stably integrated either construct were selected and expanded into clonal cell lines. Each cell line was assayed for persistent RNAi by transient cotransfection with a mixture of two reporter genes, dsRED to mark transfected cells and GFP to test for stable silencing.

Transfection of clonal P19 EC cells that had stably integrated the control vector produced equal numbers of red and green cells, as would be expected in the absence of any specific silencing response (Fig. 5*B*), whereas cells that express the GFP hairpin RNA gave a very different result. These cells expressed the dsRED protein with an efficiency comparable to that observed in cells containing the control vector. However, the cells failed to express the cotransfected GFP reporter (Fig. 5*B*). These data provide a strong indication that continuous expression of a hairpin dsRNA can provoke stable, sequence-specific silencing of a target gene.

In *Drosophila* S2 cells and *C. elegans* (18, 27–30), RNAi is initiated by the Dicer enzyme, which processes dsRNA into ≈22-nt siRNAs (18). In both, S2 cells and *C. elegans* experiments by using dsRNA to target Dicer suppress the RNAi response (18, 27, 29). Whether Dicer plays a central role in hairpin-induced gene silencing in P19 cells was tested by transfecting P19 cells stably transfected with GFP hairpin constructs with mouse *Dicer* dsRNA (see *Materials and Methods*). Treatment with *Dicer* dsRNA, but not control dsRNA, resulted in derepression of GFP (Fig. 5*C*).

dsRNA Induces Posttranscriptional Silencing. A key feature of RNAi is that it exerts its effect at the posttranscriptional level by destruction of targeted mRNAs (reviewed in ref. 4). To test

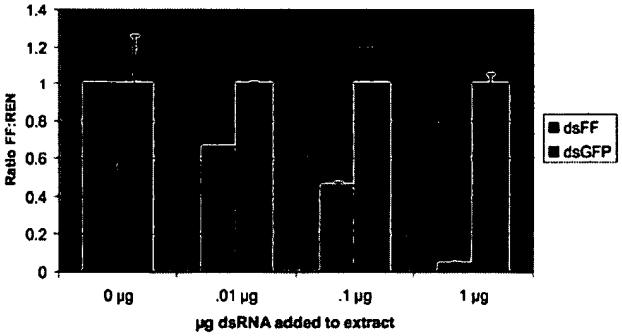


Fig. 6. dsRNA induces silencing at the posttranscriptional level. P19 cell extracts were used for *in vitro* translation of firefly and *Renilla* luciferase mRNA (100 ng each). Translation reactions were programmed with various amounts of dsRNA 500mers, either homologous to firefly luciferase mRNA (dsLUC) or nonhomologous (dsGFP). Luciferase assays were carried out after a 1-h incubation at 30°C. Ratios of firefly to *Renilla* activity are normalized to no dsRNA controls. Standard deviations from the mean are shown.

whether dsRNAs induced silencing in mouse cells via posttranscriptional mechanisms, we used an assay identical to that, used initially to characterize RNAi responses in *Drosophila* embryo extracts (25). We prepared lysates from P19 EC cells that were competent for *in vitro* translation of capped mRNAs corresponding to *Renilla* and firefly luciferase. Addition of nonspecific dsRNAs to these extracts had no substantial effect on either the absolute amount of luciferase expression or on the ratio of firefly to *Renilla* luciferase (Fig. 6). In contrast, addition of dsRNA homologous to the firefly luciferase induced a dramatic and dose-dependent suppression of activity. Addition of RNA corresponding to only the antisense strand of the dsRNA had little effect, comparable to a nonspecific dsRNA control, and pretreatment of the dsRNA silencing trigger with RNase III greatly reduced its potential to induce silencing *in vitro*. A second hallmark of RNAi is the production of small, ≈22-nt siRNAs, which determine the specificity of silencing. We found that such RNA species were generated from dsRNA in P19 cell extracts (Fig. 5*D*, *in vitro*), indicative of the presence of a mouse Dicer activity. These species were also produced in cells that stably express GFP hairpin RNAs (Fig. 5*D*, *in vivo*). Considered together, the posttranscriptional nature of dsRNA-induced silencing, the association of silencing with the production of ≈22-nt siRNAs, and the dependence of this response on Dicer, a key player in the RNAi pathway, strongly suggests that dsRNA suppresses gene expression in murine cells via a conventional RNAi mechanism.

Discussion

The discovery that dsRNA could induce gene silencing in organisms as diverse as plants and parasitic protozoans has raised the possibility that RNAi might be a nearly universal mechanism of gene silencing. This notion has been supported by the identification of homologs of proteins that participate in the silencing process in virtually all genomes examined to date, with the exception of *Saccharomyces cerevisiae* (reviewed in ref. 4). The first indications that this response might also extend to mammals came from the observation that injection of dsRNAs into early mouse embryos induced sequence-specific silencing (23, 24). Recent work by Tuschl and colleagues (5) had shown that siRNAs can induce silencing in numerous mammalian cell lines, presumably by entering the RNAi pathway. However, both in mouse embryos and previous mammalian cell culture studies, silencing was transient.

As an extension of these pioneering studies, we have demonstrated that dsRNA can induce potent and specific gene silencing in mouse embryonic cell lines. Specifically, we have shown that silencing can be induced by long dsRNAs in mouse embryonal carcinoma cell lines, in normal mouse embryonic stem cells, and in some mouse somatic cells. There are several indications that this phenomenon might be mechanistically related to RNA interference pathways that have been characterized in plants, *C. elegans*, and *Drosophila*. First, induction of silencing requires dsRNA. Second, *in vitro* studies suggest that silencing occurs at the posttranscriptional level. Third, silencing is correlated with the appearance of ≈22-nt siRNAs homologous to the gene that is being suppressed. However, final placement of the phenomenon reported here within the pantheon of dsRNA-induced silencing mechanisms will require a characterization of the protein and/or ribonucleoprotein machinery, which enforces suppression. A significant step toward this goal has been taken by the demonstration that Dicer is required for dsRNA-induced silencing in P19 cells.

We have demonstrated that stable, sequence-specific silencing can be induced by enforcing endogenous expression of RNA hairpins. The ability to create permanent cell lines with a desired loss-of-function phenotype extends the utility of RNAi as method for probing gene function in mammalian cells. This capability enables the production of large numbers of silenced cells for biochemical analysis and permits the evaluation of phenotypes over long time spans. However, perhaps the two most important ramifications of stable RNAi are the ability to harness this technology for unbiased, phenotype-based genetic selections and the possibility that stably silenced, embryonic cell lines might ultimately be used to reconstitute animals containing a specifically silenced locus.

We thank Scott Hammond and Emily Bernstein for critical reading of the manuscript. Large quantities of purified RNase III were a kind gift of Al Nicholson (Wayne State University). P.J.P. is an Arnold and Mabel Beckman Anderson Fellow of the Watson School of Biological Sciences and thanks Richard M. Paddison for academic support. A.A.C. is a George A. and Marjorie H. Anderson Fellow of the Watson School of Biological Sciences and is a predoctoral fellow of the Howard Hughes Medical Institute. G.J.H. is a Rita Allen Foundation scholar. This work was supported in part by National Institutes of Health Grant RO1-GM62534 and R21-CA91069 (to G.J.H.) and by a grant from Genetica, Inc. (Cambridge, MA).

1. Edgar, R. S. & Leilausis, I. (1964) *Genetics* **49**, 649–662.
2. Hartwell, L. H., Culotti, J., Pringle, J. R. & Reid, B. J. (1974) *Science* **183**, 46–51.
3. Hartwell, L. H. & Weinert, T. A. (1989) *Science* **246**, 629–634.
4. Hammond, S. M., Caudy, A. A. & Hannon, G. J. (2001) *Nat. Rev. Genet.* **2**, 110–119.
5. Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K. & Tuschl, T. (2001) *Nature (London)* **411**, 494–498.
6. Hamilton, A. J. & Baulcombe, D. C. (1999) *Science* **286**, 950–952.
7. Hammond, S. M., Bernstein, E., Beach, D. & Hannon, G. J. (2000) *Nature (London)* **404**, 293–296.
8. Zamore, P. D., Tuschl, T., Sharp, P. A. & Bartel, D. P. (2000) *Cell* **101**, 25–33.
9. Fraser, A. G., Kamath, R. S., Zipperlen, P., Martinez-Campos, M., Sohrmann, M. & Ahringer, J. (2000) *Nature (London)* **408**, 325–330.
10. Gonczy, P., Echeverri, G., Oegema, K., Coulson, A., Jones, S. J., Copley, R. R., Duperon, J., Oegema, J., Brehm, M., Cassin, E., *et al.* (2000) *Nature (London)* **408**, 331–336.
11. Kennerdell, J. R. & Carthew, R. W. (1998) *Cell* **95**, 1017–1026.
12. Clemens, J. C., Worby, C. A., Simonson-Leff, N., Muda, M., Maehama, T., Hemmings, B. A. & Dixon, J. E. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 6499–6503.
13. Kennerdell, J. R. & Carthew, R. W. (2000) *Nat. Biotechnol.* **18**, 896–898.
14. Williams, B. R. (1997) *Biochem. Soc. Trans.* **25**, 509–513.
15. Gil, J. & Esteban, M. (2000) *Apoptosis* **5**, 107–114.
16. Clarke, P. A. & Mathews, M. B. (1995) *RNA* **1**, 7–20.
17. Baglioni, C. & Nilsen, T. W. (1983) *Interferon* **5**, 23–42.
18. Bernstein, E., Caudy, A. A., Hammond, S. M. & Hannon, G. J. (2001) *Nature (London)* **409**, 363–366.
19. Smith, N. A., Singh, S. P., Wang, M. B., Stoutjesdijk, P. A., Green, A. G. & Waterhouse, P. M. (2000) *Nature (London)* **407**, 319–320.
20. Tavernarakis, N., Wang, S. L., Dorovkov, M., Ryazanov, A. & Driscoll, M. (2000) *Nat. Genet.* **24**, 180–183.
21. Kalejta, R. F., Brideau, A. D., Banfield, B. W. & Beavis, A. J. (1999) *Exp. Cell Res.* **248**, 322–328.
22. Connelly, J. C. & Leach, D. R. (1996) *Genes Cells* **1**, 285–291.
23. Wianny, F. & Zernicka-Goetz, M. (2000) *Nat. Cell Biol.* **2**, 70–75.
24. Svoboda, P., Stein, P., Hayashi, H. & Schultz, R. M. (2000) *Development (Cambridge, U.K.)* **127**, 4147–4156.
25. Tuschl, T., Zamore, P. D., Lehmann, R., Bartel, D. P. & Sharp, P. A. (1999) *Genes Dev.* **13**, 3191–3197.
26. Kawagishi-Kobayashi, M., Cao, C., Lu, J., Ozato, K. & Dever, T. E. (2000) *Virology* **276**, 424–434.
27. Grishok, A., Pasquinelli, A. E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D. L., Fire, A., Ruvkun, G. & Mello, C. C. (2001) *Cell* **106**, 23–34.
28. Hutvagner, G., McLachlan, J., Pasquinelli, A. E., Balint, E., Tuschl, T. & Zamore, P. D. (2001) *Science* **293**, 834–838.
29. Ketting, R. F., Fischer, S. E., Bernstein, E., Sijen, T., Hannon, G. J. & Plasterk, R. H. (2001) *Genes Dev.* **15**, 2654–2659.
30. Knight, S. W. & Bass, B. L. (2001) *Science* **293**, 2269–2271.

RNA interference microarrays: High-throughput loss-of-function genetics in mammalian cells

Jose M. Silva, Hana Mizuno, Amy Brady, Robert Lucito, and Gregory J. Hannon*

Watson School of Biological Sciences, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724

Edited by Stanley Fields, University of Washington, Seattle, WA, and approved March 3, 2004 (received for review January 9, 2004)

RNA interference (RNAi) is a biological process in which a double-stranded RNA directs the silencing of target genes in a sequence-specific manner. Exogenously delivered or endogenously encoded double-stranded RNAs can enter the RNAi pathway and guide the suppression of transgenes and cellular genes. This technique has emerged as a powerful tool for reverse genetic studies aimed toward the elucidation of gene function in numerous biological models. Two approaches, the use of small interfering RNAs and short hairpin RNAs (shRNAs), have been developed to permit the application of RNAi technology in mammalian cells. Here we describe the use of a shRNA-based live-cell microarray that allows simple, low-cost, high-throughput screening of phenotypes caused by the silencing of specific endogenous genes. This approach is a variation of "reverse transfection" in which mammalian cells are cultured on a microarray slide spotted with different shRNAs in a transfection carrier. Individual cell clusters become transfected with a defined shRNA that directs the inhibition of a particular gene of interest, potentially producing a specific phenotype. We have validated this approach by targeting genes involved in cytokinesis and proteasome-mediated proteolysis.

RNA interference (RNAi) has emerged as one of the standard techniques to study gene function in diverse experimental systems. Introduction of double-stranded RNA (dsRNA) into a cell decreases the level of the complementary mRNAs producing a knockdown of the corresponding protein. The current model of the RNAi mechanism proposes that the silencing "trigger" is processed by Dicer into small RNAs of 21–22 nucleotides in length. These become incorporated into an RNA-induced silencing complex with endonuclease activity (RISC), which, in turn, identifies and cleaves homologous mRNAs (1, 2).

Based on this approach, genomewide RNAi approaches have been used successfully for phenotype-based screens in *Caenorhabditis elegans* (3–5) and *Drosophila melanogaster* (6, 7). In part, these successes derive from the availability of convenient and inexpensive methods for producing and introducing dsRNA. For example, it has previously been shown that RNAi can be triggered by soaking *C. elegans* in a solution of dsRNA (8), or by feeding worms with *E. coli* expressing gene-specific dsRNAs (9). In *Drosophila* cells a soaking protocol is also available allowing an easy method of introducing dsRNA (10).

Unfortunately, similarly straightforward approaches for triggering silencing have not been described in mammals. Analysis of multiples genes requires a "gene by gene" method, in which individual transfections must be performed, making these studies expensive, tedious, and dependent on high-throughput robotic systems. Cell microarrays represent a novel alternative to classical approaches to phenotype-based assays in mammalian cells.

Cell microarrays were first described by Ziauddin and Sabatini (11), who demonstrated that cells grown on a glass substrate could take up DNA–lipid complexes that had been deposited on the slide before cells were plated. Cells essentially became transfected *in situ*, with defined spots of transfected cells localized over the printed DNAs. These studies demonstrated the use of conventional DNA constructs for creating phenotypes based on ectopic expression. Here we investigate the possibility of similarly using cell microarrays for loss-of-function genetics. This is accomplished by creating a microarray of living cells that have been transfected *in situ* with either small interfering RNAs (siRNAs) or with DNA constructs that direct the expression of short hairpin RNAs (shRNAs). These are effective at initiating a silencing response and in creating defined areas (spots) of cells in which suppression of a targeted gene generates an expected phenotype. Such arrays will find broad application to high-throughput low-cost phenotype-based screens in mammalian cells.

Materials and Methods

Microarray Printing and Reverse Transfection. Transfection mixes containing DNA reporter vectors (500 ng) plus shRNAs (1 μg) or siRNAs (200 ng) were printed onto glass slides by using a previously described "lipid method" (11) with some modifications. Briefly, nucleic acids were resuspended in 15 μl of DNA-condensation buffer (Buffer EC, Qiagen, Valencia, CA) with a final concentration of 0.4 M sucrose. After two incubation steps with the enhancer solution and the Effectene transfection reagent (Qiagen), a 1× volume of 0.2% gelatin was remixed with the solution to complete a transfection master mix of 45 μl. Ten microliters of this was aliquotted into a 384-well plate for printing, and the remaining 35 μl was stored at 4°C for later assays. Samples were printed onto Corning GAPS II slides with a PixSys 5500 Robotic Arrayer (Genomic Solutions, Ann Arbor, MI). Pins transferred the "lipid–DNA" solution to the slide while touching the surface of the slide for 500 ns. To ensure enough printed area to contain several hundred cells, we printed in close proximity nine spots forming a 3 × 3 square. After printing, the nine spots fuse together forming a bigger single dot with a diameter of ≈400–500 μm. To prevent contamination after printing, the slides were dried overnight at room temperature in a tissue culture hood.

To perform the reverse transfection, one spotted array was placed inside a 10-cm tissue culture dish and 15 ml of media containing cells at a concentration of 1×10^6 per milliliter was added into the dish taking care not to disturb the printed surface. Cells were incubated for 60 h without media change before analysis of the results (Fig. 1).

Reporter Assays. One hundred sixty dots containing a dual reporter vector expressing GFP/dsRed fluorescent proteins (gift of Alla Karpova, Cold Spring Harbor Laboratory) and individual shRNAs were printed. All shRNA were part of a library of U6 polymerase III promoter-driven hairpins (28). Four groups of experiments with 40 dots (each) were printed: the first group contained only dual reporter vector, the second group contained the reporter vector plus an shRNA or siRNA against firefly luciferase (Ff shRNA and Ff siRNA), the third group contained

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: Ff, firefly luciferase; RNAi, RNA interference; shRNA, short hairpin RNA; siRNA, small interfering RNA.

*To whom correspondence should be addressed. E-mail: hannon@cshl.edu.

www.pnas.org/cgi/doi/10.1073/pnas.0400165101

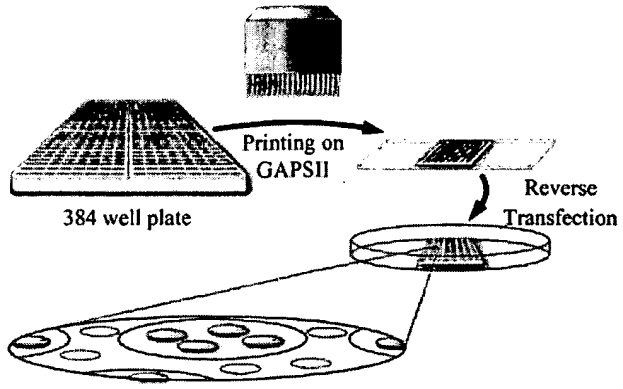


Fig. 1. Outline of the protocol used to perform reverse transfection on a glass slide. Transfection mixes containing gelatin were spotted on GAPS II glass slides with a PixSys 5500 Robotic Arrayer to form a 3 × 3 square. Dried slides were put inside a 10-cm tissue culture dish, and medium with cells was added to the dish. After incubation, groups of transfected cells could be detected inside the spotted transfection mixes.

the reporter vector plus a shRNA or a siRNA against GFP that has no effect in the expression level of the protein (GFP shRNA-1 and GFP siRNA-1), and the last group contained the reporter vector plus a shRNA that reduces by 90% the GFP signal when tested in culture plates (GFP shRNA-2 and GFP siRNA-2).

Several cell lines were tested for transfection, NIH 3T3, IMR90/E1A, HeLa, and HEK 293T. To test the stability of the printed array, we repeated the assay at different time points after printing, day 0, 1 week, 2 weeks, 4 weeks, and 2 months. For testing the stability of the transfection master mix, we stored the solution at 4°C and then printed new slides and assayed them at the time points described above.

Proteasome-Mediated Proteolysis Assays. Thirty shRNAs targeting different proteasome subunits were printed in triplicate. Every dot harbored an shRNA-expression vector, a plasmid expressing dsRed (dsRed N-1, Clontech), and a vector encoding a proteasome fluorescent reporter (ZsProSensor, Clontech). This reporter encodes a fusion protein that has been engineered to show varying levels of expression depending on the status of the proteasome pathway. Every transfection master mix contained 400 ng of dsRed vector, 100 ng of ZsProSensor, and 1 μg of shRNA plasmid. Twenty micrograms of total protein lysates was used for Western blot analysis. Rabbit anti-PSMC-6 subunit of the proteasome (Affinity, Biomol, Plymouth Meeting, PA), rabbit anti-ubiquitin (StressGen Biotechnologies, Victoria, Canada), and mouse anti-β-actin (United States Biological, Swampscott, MA) antibodies were also used in these studies.

Cytokinesis Defect Assays. Eight shRNAs targeting the motor protein Eg5 were printed (10 replicas each) together with a vector encoding an α-tubulin-GFP fusion protein (GFP-tubulin, Clontech). Every transfection mix contained 1 μg of Eg5 hairpin and 500 ng of the fluorescent fusion protein expression plasmids.

For immunofluorescence studies, we printed a replica slide where the transfection mix contained 500 ng of the dsRed reporter instead of the GFP-tubulin. Cells were stained by using standard methods with small variations. After incubation of the slides for 60 h the cells were fixed with paraformaldehyde for 10 min, washed very gently, and permeabilized with 1% Triton X-100 in PBS for 15 min on ice. Only one 15-min wash was performed to avoid washing away the cells. Mouse anti-α-tubulin (Sigma) was used in this study. Hoechst dye was included in the last wash to visualize the chromosomes.

Ninety-Six-Well Plate Analyses. All RNAi microarray results were validated by using cells transfected in 96-well tissue culture plates. Cells were transfected with LT-1 (Mirus, Madison, WI) according to the manufacturer's instructions at 50–70% confluence. The plasmids containing appropriate constructs were cotransfected, keeping the same ratios used in the arrayed slides but with a total mass of 100 ng of DNA for each transfected well. Again, results were analyzed after 60 h of incubation.

Results

Targeting Reporter Genes *in Situ* by Using siRNAs. Given previous successes in ectopically expressing genes by reverse transfection (11), we hoped that similar approaches could be coupled with the use of RNAi to produce knockdown phenotypes. Therefore, we began by testing the ability of siRNAs to be deposited on a microarray as lipid–RNA complexes and to cause sequence-specific silencing in cells grown on the arrays. We began by testing the ability of siRNAs to silence a co-delivered, ectopic marker. For convenience, we used GFP to enable both transfection and silencing to be scored by visual inspection. Because GFP was the siRNA target, we also included a plasmid that directs the expression of a second fluorescent protein (dsRed) to allow us to verify that siRNAs specifically silenced GFP expression rather than simply interfered with transfection.

Several siRNAs homologous to the GFP coding sequence were mixed with plasmids encoding GFP and dsRed. Nucleic acids were combined with a variety of lipid reagents [LT-1 or TKO (Mirus) and Effectene (Qiagen)], and the lipid–nucleic acid complexes were spotted onto glass slides. We found that Qiagen reagent performed best, giving optimal transfection of both DNA and RNA in this experiment. GFP siRNAs that had been previously identified as an efficient suppressor of the expression of the fluorescent protein in standard transfections (GFP siRNA-2) also showed potency on the spotted array. Ineffective siRNAs (GFP siRNA-1) or unrelated siRNAs (Ff siRNA) did not produce any effect on GFP expression levels either in standard transfections or on microarray slides (Fig. 2*A*).

Targeting Reporter Genes *in Situ* by Using shRNAs. The aforementioned studies demonstrated that an RNAi response could be initiated by *in situ* transfection of siRNAs on glass-slide microarrays. As stated above, RNAi can also be initiated in mammalian cells by transfection of DNA expression constructs that direct the synthesis of shRNA sequences. We therefore tested the possibility that a slide printed with a transfection mix containing an shRNA-expression construct that targets a specific mRNA could initiate a specific silencing response in *in situ* transfected cell clusters. As a first step, we again tested our ability to down-regulate the expression of GFP. Spots were printed as described above. In every cell cluster, the GFP level reported specific RNAi-mediated suppression and the dsRed level again acted as a transfection control.

As expected, none of the spots containing the reporter vector alone, the reporter plus a control shRNA, targeting firefly luciferase, or the reporter plus GFP shRNA-1, an ineffective shRNA, showed any reduction in the expression of GFP. In contrast, all spots harboring the GFP shRNA-2, an active shRNA, presented a strong attenuation of the GFP signal while maintaining unaltered levels of dsRed protein (Fig. 2*B*). It is worth noting that all suppressed samples showed a similar degree of attenuation of GFP and a very similar percentage of transfected cells, revealing a high degree of consistency from spot to spot.

We next asked whether the *in situ* RNAi procedure could be applied broadly by examining the response of a number of different cell lines in an arrayed format. These included transformed and nontransformed cells lines, fibroblasts and epithelial cells, and lines from mouse and human. Of those tested, HEK

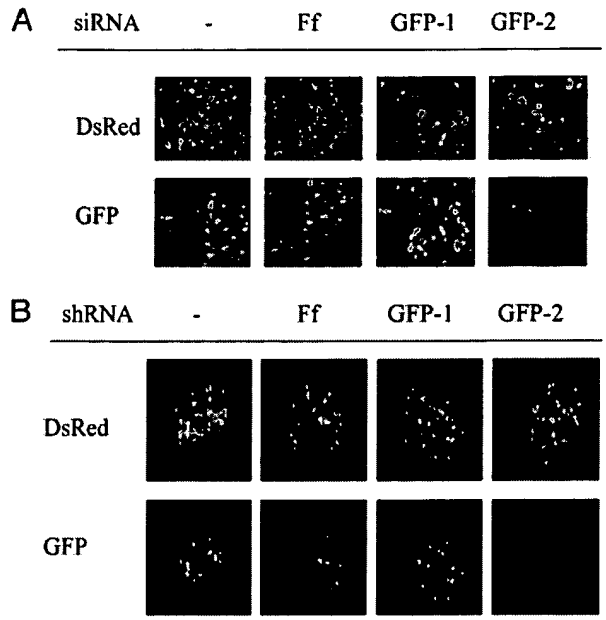


Fig. 2. The GFP reporter is specifically suppressed by RNAi in 293T cells incubated for 48 h on a printed glass slide. (*A*) Printed spots contained a vector expressing GFP and dsRed reporters and individual siRNAs targeting different sequences (as indicated): no siRNA (—), Ff, siRNA-1 (GFP-1), and siRNA-2 (GFP-2). (*B*) Printed spots contained a vector expressing GFP and dsRed reporters and individual shRNA expression plasmids targeting different sequences (as indicated): no hairpin (-), Ff, hairpin 1 (GFP-1), and hairpin 2 (GFP-2). Control spots harbored siRNA or an shRNA-expression vector against Ff, GFP-1, or no added RNAi inducer (mock).

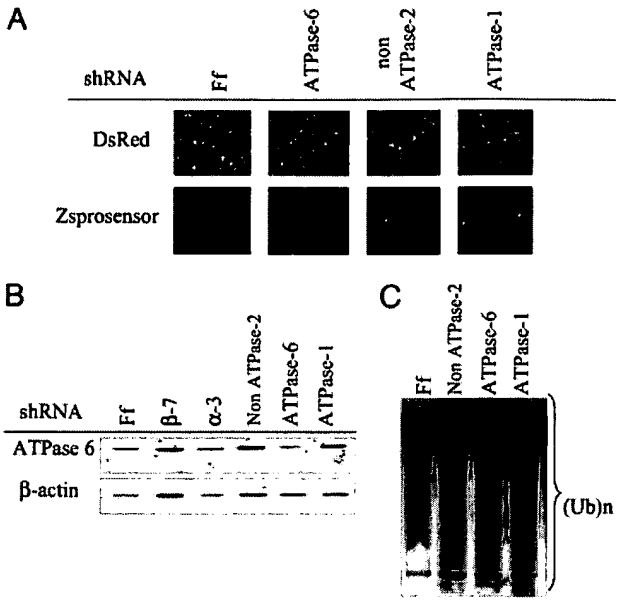


Fig. 3. (*A*) Different levels of ZsGreen protein accumulation detected in the RNAi microarray. (*B*) HEK 293T were incubated with individual hairpins targeting the expression of different proteasome subunits. A Western blot shows specific inhibition of ATPase-6 shRNA. Firefly shRNA was used as control. (*C*) Anti-ubiquitin Western blot showing accumulation of polyubiquitinated proteins in treated cells.

293T cells showed the highest efficiency. IMR90/E1A, NIH 3T3, and HeLa cells showed lower efficiency (20–50%) than 293T (data not shown). Our results indicate that by varying the lipid and nucleic acid content of the spot, many different cell lines can be used. However, for convenience, we performed the remainder of our assays with HEK 293T.

We also examined the stability of slides spotted with lipid/shRNA mixtures and with lipid/siRNA mixtures. We printed replica slides and compared both transfection efficiency and silencing efficiency at various time points. We did not observe any reduction in the performance of the shRNA arrays after 1 week, 1 month, or up to 2 months of storage at 4°C. In contrast, siRNA printed slides showed more variability, and unclear results were obtained after 2 weeks of storage.

An *in Situ* Assay for Defects in Proteosome Function. The need to locate transfected cells to score phenotypes on the array led to the consideration of phenotypic assays that depended on the expression of a fluorescent, exogenous reporter gene. The basic idea, a variation of the validation experiments described above, was to include in the transfection mix a plasmid harboring a fluorescent reporter that is differentially expressed (at higher or lower level) in cells in which RNAi has been used to alter the function of a specific pathway. To test this approach, we focused on protein half-life as a determinant of steady-state expression levels. A variety of motifs have been found to confer a short lifetime on cellular proteins. So-called PEST sequences are rich in the amino acids Pro (P), Asp, Glu (E), Ser (S), and Thr (T), which occur in internal positions in the sequence. Proteins containing PEST sequence elements are rapidly targeted to the 26S proteasome for degradation (12).

The mouse ornithine decarboxylase (MODC) has an extremely short half-life (13). MODC degradation is mediated by an internal domain, called MODC-d410, that contains several PEST motifs. This functional motif is transferable, decreasing greatly the stability of proteins to which it is fused (14, 15). This property has been exploited to create commercially available reporter systems in which MODC-d410, appended to a fluorescent protein, creates a fusion that can indicate the integrity of the proteosome in living cells.

We chose to analyze the expression level of a commercially available reporter consisting of a green fluorescent protein (ZsGreen) tagged on the carboxyl-terminus with the MODC-d410 domain (ZsProSensor, Clontech). We recently screened a library of 7,000 individual shRNAs for the ability to antagonize proteosome function in a 96-well plate format (28). Roughly one-quarter of primary positives targeted proteosome subunits. Putative positives from that screen were used on microarrays to validate the *in situ* approach and to compare the sensitivity of array-based assays to those carried out in 96-well plates. Based on these previous studies, thirty different hairpins targeting proteasome subunits were deposited, each in individual spots, together with the destabilized green reporter and a dsRed vector. As in previously described experiments, dsRed served as transfection control.

After 24 h of incubation we observed higher levels of ZsGreen protein in several dots containing proteasome shRNAs compared to control shRNAs, whereas no changes were observed in the intensity of the red fluorescent protein. During the following 48 h the green fluorescent signal gradually increased in these positive spots, achieving maximum intensity after 60–72 h. We identified clear differences in the accumulation level of the reporter in spots containing proteasome shRNAs compared with nonproteasome hairpins. Interestingly, the most intense signals were revealed from spots containing shRNAs that target subunits of the 19S base (Fig. 3*A*).

To confirm that the increased expression of the reporter was associated with alterations in proteasome function, we verified that cells transfected with a hairpin against the ATPase subunit 6 indeed showed a specific reduction of the targeted protein. As expected, Western blot analysis of transiently transfected 293T

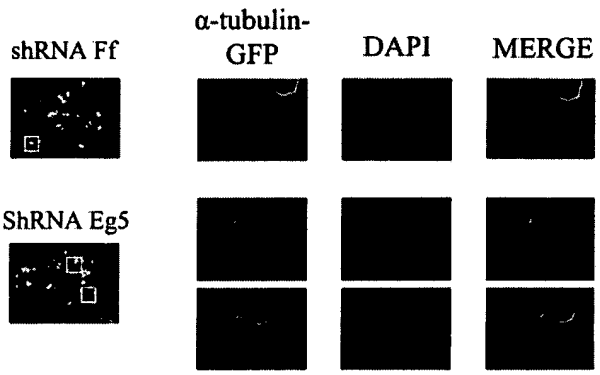


Fig. 4. Cytokinesis defects induced by Eg5 shRNAs. (*Upper*) Sample of normal mitotic metaphase detected in a dot containing firefly (control) shRNA. (*Lower*) Two typical "rosette" phenotype found in a printed spot containing the shRNA-7 against the motor protein Eg5.

cells with the ATPase-6 shRNA-1 showed a specific reduction in the level of the targeted protein (Fig. 3*B*).

It has been shown previously that in cells in which the normal function of the proteasome is blocked, there is an accumulation of polyubiquitinated proteins (16, 17). To validate further the antagonism of proteosome function by RNAi, we examined the level of polyubiquitinated proteins. Indeed, analysis of bulk ubiquitinated proteins by Western blotting with a ubiquitin antibody revealed an increase in these normally unstable species in cells treated with proteasome shRNAs compared with controls (Fig. 3*C*).

A Screen for Alterations in Cell Cycle Control: Cytokinesis Defects. The aforementioned screen was amenable to the use of microarray scanners to score positives by their ratio of fluorescent signals. To examine the suitability of this approach for other types of screens, we performed a live-cell microarray assay to identify cytokinesis defects induced by RNAi. As a reconstruction experiment, we knocked down the mitotic motor protein Eg5, because cytokinesis defects in cells where Eg5 function is inhibited are well established (18, 19).

The mitotic spindle, which consists of a dynamic array of microtubules and associated proteins, is responsible for segregation of chromosomes during mitosis. Studies using immunodepletion (20, 21) of the motor kinesin Eg5 or with the specific inhibitor monastrol (18) have revealed that inhibition of the normal function of this protein causes a defect in spindle formation. Initially, a defect in centrosome separation causes the assembly of monopolar spindles, and eventually, aberrant structures form that look like "rosettes" of microtubules with DNA at the periphery.

Based on the expectation of this typical morphology, we printed DNA–lipid complexes containing individual shRNAs targeting Eg5 on microarray slides. Transfection mixes also contained a plasmid encoding an α-tubulin GFP fusion protein. In this array, the GFP fusion protein identifies the cells that have been transfected and allows visualization of microtubules. Additionally, when the microarray was scored, Hoechst 33342 was added to the media to allow visualization of chromosomes. The analysis of the array revealed two hairpins that produced cells displaying the characteristic "rosette" pattern (Fig. 4). We did not observe this phenotype in any of the spots containing control shRNAs. This phenotype was also observed when shRNAs targeting Eg5 were tested in a 96-well plate format and was similar to the phenotype obtained after Eg5 inhibition by monastrol (data not shown).

The use of vectors harboring reporters or fusion proteins is a very convenient approach to identify abnormal phenotypes by *in situ* transfection. However, appropriate reagents will not be easily available for all interesting phenotypes. For this reason, we asked whether it was possible to identify Eg5 suppression phenotypes by the conventional technique of immunofluorescence (IF). A replica of the Eg5 glass slide, in which the α-tubulin fluorescence reporter was replaced by dsRed, was stained with a standard IF protocol. We find that we could easily identify cytokinesis defects produced by the same shRNAs identified with the α-tubulin GFP protein by using this methodology (data not shown).

Discussion

Genomewide analyses of loss-of-function phenotypes in mammals, similar to classic genetic studies in yeast, were very difficult, if not impossible, only a few years ago. However, over the last several years RNAi has emerged as a powerful approach for manipulating gene expression in mammalian cells, opening the door to the execution of such screens. High-throughput RNAi analyses have previously been used to study gene function in *D. melanogaster* and *C. elegans*. For example, essential genes (22), G protein-coupled receptors (GPCRs) (23), fat regulatory genes (24), and genes that regulate lifespan (25) have been functionally analyzed by this approach. Unfortunately, similar studies in mammals still represent a technological challenge. In part, limitations occur because of the cost of the RNA species themselves and the cost of introducing these species (siRNA and shRNA) into mammalian cells. Additionally, the use of a large-scale analysis in any system is limited by the screening methodology.

Here we validate the use of a cost-effective high-throughput procedure for RNAi-based screens in mammalian cells. This procedure is based on methodologies developed by Sabatini and colleagues (11) for creating high copy suppression phenotypes by "reverse transfection." This allows for the cost-effective use of materials, both the nucleic acids themselves and tissue culture reagents. We estimate that between 100 and 500 reverse transfections can be done with the materials required for a single transfection in a well of a 96-well plate. Additionally, thousands of samples can be printed in parallel on a glass-slide microarray, reducing the time and cost associated with maintaining cultures and analyzing phenotypic outputs. Finally, as previously described, printed slides can be stored for several months without losing potency (11).

By comparing results obtained by initiating RNAi *in situ* on microarrays to screens conducted in 96-well plates, we find that the arrays compare favorably to standard methods for both sensitivity and specificity. In agreement with our data, two recent papers have reported that RNAi could be initiated on printed slides to inhibit the expression of a co-delivered marker gene (26, 27). Our study extends these results by showing that a similar procedure can be used to silence endogenous genes to create RNAi-induced phenotypes relevant to two independent biological pathways.

Firstly, we designed an assay to study proteosome-mediated degradation. As predicted, when subunits of the 26S proteasome were targeted by shRNAs, we could clearly identify accumulation of an engineered protein that is normally degraded by the proteasome pathway. We showed that proteolytic defects were due to specific suppression of the targeted 26S subunits and demonstrated that cellular levels of the natural proteasome substrates (ubiquitinated proteins) were affected in the same way as the fluorescent reporter. In a second study, we analyzed the effect produced by loss of kinesin Eg5 as model for cytokinesis defects. After targeting with Eg5 shRNAs, we could reproduce the expected aberrant spindle morphology in a printed slide format, while no changes were observed in control spots.

Here, we demonstrate the feasibility of using printed arrays of siRNAs and shRNAs for highly parallel phenotype analysis in living cells. This approach is flexible and provides low-cost alternatives to similar screens carried out in multiwell plate formats. As large libraries of shRNAs become widely available

(28, 29), the techniques described herein will become a powerful approach to genetic analysis in mammalian cells.

We thank Jim Duffy for help with preparation of the figures and Michelle Carmell, Ahmet Denli, Liz Murchinson, Lin He, and William M. Keyes for critical reading of the manuscript. J.M.S. is supported by a postdoctoral fellowship from the U.S. Army Prostate Cancer Research Program, and G.J.H. is supported by an Innovator Award from the U.S. Army Breast Cancer Research Program. This work was supported in part by grants from the National Institutes of Health.

1. Zamore, P. D. (2001) *Nat. Struct. Biol.* **8,** 746–750.
2. Hannon, G. (2002) *Nature* **418,** 244–251.
3. Maeda, I., Kohara, Y., Yamamoto, M. & Sugimoto, A. (2001) *Curr. Biol.* **11,** 171–176.
4. Ashrafi, K., Chang, F. Y., Watts, J. L., Fraser, A. G., Kamath, R. S., Ahringer, J. & Ruvkun, G. (2003) *Nature* **421,** 268–272.
5. Pothof, J., van Haaften, G., Thijssen, K., Kamath, R. S., Fraser, A. G., Ahringer, J., Plasterk, R. H. & Tijsterman, M. (2003) *Genes Dev.* **17,** 443–448.
6. Somma, M. P., Fasulo, B., Cenci, G., Cundari, E. & Gatti, M. (2002) *Mol. Biol. Cell.* **13,** 2448–2460.
7. Kiger, A., Baum, B., Jones, S., Jones, M., Coulson, A., Echeverri, C. & Perrimon, N. (2003) *J. Biol.* **2,** 27.
8. Tabara, H., Grishok, A. M. &. Mello, C. C. (1998) *Science* **282,** 430–431.
9. Timmons, L. & Fire, A. (1998) *Nature* **395,** 854.
10. Clemens, J. C., Worby, C. A., Simonson-Leff, N., Muda, M., Maehama, T., Hemmings, B. A. & Dixon, J. E. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 6499–6503.
11. Ziauddin, J. & Sabatini, D. M. (2001) *Nature* **411,** 107–110.
12. Rechsteiner, M. & Rogers, S. W. (1996) *Trends Biochem. Sci.* **21,** 267–271.
13. Murakami, Y., Matsufuji, S., Kameji, T., Hayashi, S., Igarashi, K., Tamura, T., Tanaka, K. & Ichihara, A. (1992) *Nature* **360,** 97–99.
14. Olmo, M. T., Rodríguez-Agudo, D., Medina, M. A. & Sanchez-Jimenez, F. (1999) *Biochem. Biophys. Res. Commun.* **257,** 269–272.
15. Verma, R. & Deshaies, R. J. (2000) *Cell* **101,** 341–344.
16. Wojcik, C. & DeMartino, G. N. (2002) *J. Biol. Chem.* **277,** 6188–6197.
17. Bochtler, M., Ditzel, L., Groll, M., Hartmann, C. & Huber, R. (1999) *Annu. Rev. Biophys. Biomol. Struct.* **28,** 295–317.
18. Mayer, T. U., Kapoor, T. M., Haggarty, S. J., King, R. W., Schreiber, S. L. & Mitchison, T. J. (1999) *Science* **286,** 971–974.
19. Goshima, G. & Vale, R. D. (2003) *Cell Biol.* **162,** 1003–1016.
20. Sawin, K. E., LeGuellec, K., Philippe, M. & Mitchison, T. J. (1992) *Nature* **359,** 540–543.
21. Blangy, A., Lane, H. A., d'Herin, P., Harper, M., Kress, M. & Nigg, E. A. (1995) *Cell* **83,** 1159–1169.
22. Maeda, I., Kohara, Y., Yamamoto, M. & Sugimoto, A. (2001) *Curr. Biol.* **11,** 171–176.
23. Keating, C. D., Kriek, N., Daniels, M., Ashcroft, N. R., Hopper, N. A., Siney, E. J., Holden-Dye, L. & Burke, J. F. (2003) *Curr. Biol.* **13,** 1715–1720.
24. Ashrafi, K., Chang, F. Y., Watts, J. L., Fraser, A. G., Kamath, R. S., Ahringer, J. & Ruvkun, G. (2003) *Nature* **421,** 268–272.
25. Lee, S. S., Lee, R. Y., Fraser, A. G., Kamath, R. S., Ahringer, J. & Ruvkun, G. (2003) *Nat. Genet.* **33,** 40–48.
26. Mousses, S., Caplen. N. J., Cornelison, R., Weaver, D., Basik, M., Hautaniemi, S., Elkahloun, A. G., Lotufo, R. A., Choudary, A., Dougherty, E. R., *et al.* (2003) *Genome Res.* **13,** 2341–2347.
27. Kumar, R., Conklin, D. S. & Mittal, V. (2003) *Genome Res.* **13,** 2333–2340.
28. Paddison, P. J., Silva, J. M., Conklin, D. S., Schlabach, M., Li, M., Aruleba, S. Balija, V., O'Shaughnessy, A., Gnoj, L., Scobie, K., *et al.* (2004) *Nature* **428,** 427–431.
29. Berns, K., Hijmans, E. M., Mullenders, J., Brummelkamp, T. R., Velds, A., Heimerikx, M., Kerkhoven, R. M., Madiredjo, M., Nijkamp, W., Weigelt, B., *et al.* (2004) *Nature* **428,** 431–437.

TECHNICAL REPORT

Second-generation shRNA libraries to the mouse and human genomes

Jose M Silva [1#], Mamie Li [2#], Ken Chang [1#], Wei Ge [3], Mike Golding [1], Ricky Rickles[2], Despina Siolas [1], Guang Hu[2], Patrick J. Paddison [1], Michael R. Schlabach[2], Nihar Sheth [1], Jeff Bradshaw[3], Julia Burchard[3], Guy Cavet[3], Ravi Sachidanandam[1], W. Richard McCombie [1], Michele A. Cleary [3], Stephen J. Elledge [2*], and Gregory J. Hannon[1*]

[1] Cold Spring Harbor Laboratory
Watson School of Biological Sciences
1 Bungtown Road
Cold Spring Harbor, NY 11724

[2] Department of Genetics
Center for Genetics and Genomics
Harvard Medical School
Room 158D, NRB
77 Avenue Louis Pasteur
Boston, MA 02115

[3] Rosetta Inpharmatics LLC,
A Wholly Owned Subsidiary of Merck & Co., Inc.
401 Terry Ave North
Seattle, Washington 98109, USA

*To whom correspondence should be addressed
#These authors contributed equally

Loss-of-function phenotypes often hold the key to understanding the connectivity and biological functions of biochemical pathways. We and others have previously constructed libraries of short hairpin RNAs (shRNAs) that allow systematic analysis of RNAi-induced phenotypes in mammalian cells. Here we report the construction and validation of second-generation shRNA (shRNA^{mir}) expression libraries that have been designed based on an increased knowledge of RNAi biochemistry. In these improved reagents, silencing triggers have been designed to mimic a natural microRNA primary transcript, and each target sequence has been selected based on thermodynamic criteria for optimal small RNA performance. Biochemical and phenotypic assays have indicated that the new libraries are substantially improved compared to first-generation reagents. We have generated large-scale arrayed, sequence verified libraries comprising more than 125,000 shRNA^{mir} expression plasmids, covering 2/3 of all predicted genes in the human and mouse genomes. These libraries are presently available to the scientific community.

Introduction

Most eukaryotic cells harbor a natural response to double-stranded RNAs (dsRNA) that inhibits gene expression in a sequence-specific manner [1]. DsRNA silencing triggers are processed into small RNAs (siRNAs and miRNAs) that engage the RNA-induced silencing complex (RISC) to suppress expression of homologous targets. In cases in which the small RNA is perfectly complementary to the target, that RNA is cleaved and ultimately destroyed [1,2]. This pathway, known as RNA interference (RNAi), has been exploited in organisms ranging from plants to fungi to animals for deciphering gene function through suppression of gene expression. Particularly in systems where targeted genetic manipulation is difficult or time consuming, RNAi has transformed the way in which gene function can be approached on a single gene or genome-wide level [3-7].

In mammals, RNAi can be initiated in several ways. First, RNA molecules can be produced chemically [8] or enzymatically *in vitro* [9-12] and delivered to a cell. The most prevalent method of triggering RNAi is the delivery of one or more small interfering RNAs (siRNAs). SiRNAs are duplexes of ~21-22 nucleotides in length that bear two nucleotide 3' overhangs. One strand of the siRNA is incorporated into the effector complex of RNAi, the RNA-induced Silencing Complex, RISC, through the action of a RISC Loading Complex, RLC. Once in RISC, the siRNA guides substrate selection via base pairing to its complementary target [1,2].

At the heart of RISC is an Argonaute protein [13], which directly contacts the siRNA [14]. When the mRNA is engaged by this complex, the siRNA-mRNA interaction places the target in the correct alignment with the nuclease active site or "slicer" within the Argonaute PIWI domain, and the target is endonucleolytically cleaved [15-17].

The RNAi machinery can also be programmed by endogenous sources of double-stranded RNA. The most well characterized source of endogenous triggers for the RNAi machinery are the microRNA genes [18,19]. It was initially assumed that microRNAs were transcribed from the genome as short, hairpin RNAs [20] that were directly processed by Dicer to yield the mature small RNAs that enter RISC [21-24]. Over the past year, however, a different picture has emerged. Numerous studies have now demonstrated that, in animals, miRNAs are transcribed by RNA polymerase II to generate long primary polyadenylylated RNAs (pri-miRNAs) [25,26]. These primary transcripts probably adopt a complex secondary structure and fold, in the areas that harbor the mature microRNA sequences, into double-stranded RNA hairpins. Through mechanisms not yet fully understood, the pri-microRNA is recognized and cleaved at a specific site by the nuclear Microprocessor complex [27-31]. This contains an RNase III family enzyme, Drosha, that cleaves the hairpin to produce a ~70-90 nucleotide microRNA precursor (pre-miRNA) with a 2 nucleotide 3' overhang [27]. This distinctive structure signals transport of the pre-miRNA to the cytoplasm by a

mechanism mediated by Exportin-5 [32,33]. Only then is the pre-miRNA recognized by Dicer and cleaved to produce a mature microRNA. This probably involves recognition the 2 nucleotide 3' overhang created by Drosha to focus Dicer cleavage at a single site ~22 nucleotides from the end of the hairpin [14,34].

Mature miRNAs are superficially symmetrical, with 2 nucleotide 3' overhangs at each end having been generated by Drosha and Dicer, respectively. However, the individual strands of the mature miRNA enter RISC in an unequal manner. As with siRNAs, the thermodynamic asymmetry of the Dicer product is sensed such that the strand with the less stable 5' end has a greater propensity to enter RISC and guide substrate selection [35,36]. This observation of thermodynamic asymmetry within small RNAs led to the development of rules for predicting effective siRNA sequences that have greatly improved the efficiency of those RNAs as genetic tools.

Previously, several groups, including our own, described the design and construction of arrayed short hairpin RNA (shRNA) libraries that covered a fraction of human genes [37,38]. At the time when these tools were developed, our knowledge of microRNA maturation was relatively incomplete. This led most groups to the notion of expressing a simple hairpin RNA that mimicked the pre-miRNA. As our knowledge of the microRNA processing pathway and our understanding of strand preferences for RISC loading have grown, it seemed prudent to reevaluate whether the performance of encoded triggers of the RNAi pathway might be improved by remodeling a primary miRNA transcript to experimentally alter its targeting capability. Indeed such strategies have previously succeeded in both plants and animals [39,40].

Here we report the construction of a new generation of shRNA libraries (shRNA^{mir}) that takes into consideration our advancing understanding of microRNA biogenesis. In these constructs, the shRNA is harbored within the backbone of the primary mir-30 microRNA. This natural configuration proved to be up to 12 times more efficient in the production of the mature synthetic miRNAs than the previous design. Additionally, we have biochemically characterized processing of these synthetic microRNAs, allowing us to predict the exact, single mature small RNA product that will be generated from each vector. This has allowed selection of target sequences that maximize efficiency by directing preferential incorporation of the correct strand into RISC. Using these criteria, we have produced and sequence-verified more than 125,000 shRNAs covering 2/3 of the predicted genes in the mouse and human genomes. We have assayed a selected subset of shRNAs from the library for their ability to knock-down the expression of targeted genes by quantitative RT-PCR. We have also tested this set in a phenotypic assay and compared the performance of the first- and second-generation library designs. Overall, the shRNA^{mir} libraries that we describe here provide a convenient, flexible and effective tool for studying gene function in human cells. Additionally, they, for the first time, extend the possibility of large-scale RNAi screens to mouse systems.

Results

Design and construction of second generation shRNA libraries

We have previously shown that expression of a simple, 29 basepair (bp) hairpin from a U6 snRNA promoter can induce effective suppression of target genes when delivered either transiently or stably from integrated constructs (Westbrook et al., in press) [37,41]. We also found that longer hairpin structures were more effective inhibitors of gene expression than were shorter structures with stems of 19-21 nucleotides. All of these constructs, however, were designed, based upon our understanding of miRNA biogenesis at that time, to express a pre-miRNA hairpin, an intermediate in microRNA biogenesis, rather than a transcript that closely resembles a primary microRNA. Cullen and colleagues had previously shown that effective suppression could be achieved by redesigning an endogenous microRNA, miR-30, such that its targeting sequence was directed against a reporter gene [39]. We sought to compare directly the abundance of small RNAs produced from vectors with simple hairpin structures to those that more closely resemble a natural microRNA. Since it had been previously shown that the efficient ectopic expression of endogenous microRNAs requires substantial flanking sequence [42], we developed a vector in which sequences from a remodelled miR30 are flanked by ~125 bases of 5' and 3' sequence derived from the primary transcript. Incorporation of appropriate cloning sites into this vector required altering only 3 positions in the precursor. This cassette was inserted into a vector equivalent to that in which we constructed our first-generation shRNA library (pSM1), with the new shRNA vector being designated, pSM2. To distinguish the second-generation shRNAs from those in our first-generation library, we have dubbed these shRNA^{mir}.

To test the performance of pSM2 in comparison to pSM1, we used both vectors to express a sequence targeting firefly luciferase. The sequence was inserted such that an identical mature small RNA would be generated from each construct following processing *in vivo* (Fig 1a). Of primary concern was the overall amount of mature small RNA that would be generated from each construct. This was critical as dose-response experiments for shRNAs indicate that suppression correlates very well with the amount of RNA delivered [34], particularly at the relatively low doses that are expected to be achieved by expression from transfected or integrated constructs as compared to directly transfected synthetic RNAs. We transfected pSM1-luc and pSM2-luc into 293 cells, prepared RNA and assayed the processed small RNA by northern blotting. Cells transfected with pSM2-luc contained roughly 12-fold more of the small RNA than did cells transfected with pSM1-luc (Fig 1b).

As it is now clear that primary microRNAs are transcribed mainly by RNA polymerase II [25,26], we wished to compare the performance of shRNA^{mir}s driven by a variety of different promoters. We therefore cloned two different shRNA^{mir}

cassettes targeting firefly luciferase downstream of three different RNA polymerase III promoters (tRNA-val [43], U6 [41] and H1 [44]) and two different RNA polymerase II promoters (MSCV-LTR and CMV [45]). These constructs were each prepared in a plasmid backbone that carried no other mammalian promoter. Each was transfected in combination with a homologous target expression plasmid encoding firefly luciferase and with a non-targeted reporter plasmid, encoding *Renilla* luciferase, as a means of normalization. We compared the performance of these plasmids in a five different cell lines including three from human (HEK-293T, MBA-MD-231, HCT-116), one from mouse (NIH-3T3) and one from dog (MDCK). When the ability of these constructs to suppress the luciferase target was compared using a very efficient shRNA^{mir} (luc1309), we saw virtually no difference in the performance of the various promoters (Fig 1c). However, when a less efficient shRNA^{mir} (luc311) was used, differences became apparent (Fig. 1c). In this, and numerous experiments with other shRNAs (not shown), the U6 snRNA and CMV promoters gave the best and most consistent repression. The MSCV LTR, tRNA val and H1 promoters worked less efficiently overall. Based upon these studies, we chose to retain the U6 snRNA promoter in our base library vector, pSM2. It is important to note that all of our studies have been carried out in transient assays. In situations in which constructs are stably integrated into the genome at single copy, different configurations of promoters and flanking sequences perform more efficiently than U6 (see accompanying paper by Dickins et al., and Stegmeier et al., submitted).

Based upon these tests we constructed our second-generation shRNA library vector, pSM2, as shown in figure 2a. The shRNA^{mir} expression cassette is carried within a self-inactivating murine stem cell virus. Expression of the small RNA is driven by the U6 snRNA promoter. As with the first-generation shRNAs, a U6 snRNA leader sequence lies between the promoter and the 5' end of the miR-30 flanking region. Synthetic oligonucleotides encoding shRNAs are inserted into XhoI and EcoRI sites that lie within the miR-30 primary microRNA sequences. Immediately following the miR-30 cassette in each vector is a RNA polymerase III termination signal and a randomly generated 60 nucleotide barcode region to facilitate tracking of individual hairpin RNAs in complex populations. This feature is similar to that described for our first-generation RNAi library. The pSM2 vector is also designed such that inserts can be moved by an *in vivo* recombination strategy (MAGIC) [46]. Key elements of this feature are the presence on the plasmid backbone of a protein-dependent origin of replication, RK6γ and a transfer origin (oriT) that is dependent upon a complementing locus in the host cells. To permit recombination into the recipient plasmid, the shRNA^{mir} cassette is flanked by I-SceI restriction sites which, when cut in the recipient strain, reveal homology regions for recombination into the recipient plasmid. One key difference between in the second-generation shRNA libraries is that the 5' homology region is the miR-30 flanking sequence itself rather than an artificial sequence. Thus, in the second generation libraries, the shRNA cassette is transferred without the U6 snRNA promoter. This allows the construction of mating recipients that contain inducible or tissue specific

promoters (Stegmeier et al., submitted). Finally, the pSM2 vector can be selected for integration into target cells using a puromycin selection marker.

Six different shRNA^{mir} sequences were designed for each of 34,711 different known and predicted human genes and 32,628 mouse genes. In each case, shRNAs were designed such that the mature small RNA generated from each construct followed thermodynamic asymmetry rules that have been successfully applied for the design of siRNAs. This was possible because through a series of *in vitro* and *in vivo* primer extension experiments, we have mapped the precise cleavage sites used by Dicer and Drosha in transcripts generated from the pSM2 vector. This enabled us to predict precisely the small RNA that will be generated from our primary transcript (not shown).

Construction of the library proceeded stochastically using a highly parallel *in situ* synthesis approach for oligonucleotide production (Figure 2b). Groups of ~22,000 oligonucleotides, each containing a different shRNA^{mir} cassette were synthesized on glass-slide microarrays [47]. Populations were eluted from the arrays and amplified by PCR. In order to insure efficient cloning, the pSM2 backbone was inserted into a lambda phage backbone such that it was flanked by loxP sites. λ-pSM2 contains unique XhoI, EcoRI for subcloning amplified hairpins and unique FseI and AvrII sites for insertion of bar code 60mers. λ-pSM2 was first barcoded with a mixed library of random 60 nucleotide sequences amplified with a primer set which included one primer with an FseI site and one primer with a T7 promoter followed by the AvrII site. Amplified barcoded λ-pSM2 libraries were lysogenized into a strain we constructed for this purpose, DH10βλ_{KP}, which has a wild-type *pir1* gene and the lambda repressor, *cI*, to allow λ-pSM2 to replicate as a 42 kb plasmid. Approximately 10^8 $Cm^R Km^R$ lysogens were selected and served as a bar coded library pool. Bar coded λ-pSM2 was CsCl purified, then cleaved with EcoRI and XhoI before being ligated to gel purified EcoR1-XhoI cleaved pooled shRNA^{mir} inserts from an individual chip and packaged. Average library sizes were ~$5x10^7$. To generate pSM2 library plasmid pools, the phage were used to infect an *E. coli* strain we constructed, BUN25, that expresses both Cre recombinase and the *pir1-116* gene, needed for high copy RK6γ replication. Pooled plasmid libraries were then transformed into a mating competent host strain (BW F'DOT) and individual clones were sequenced at random. Clones with perfect inserts represented between 25 and 50% of the population, and these were selected and saved as an arrayed set. Accumulation of new clones from each pool was monitored dynamically and once a pool began to yield fewer unique clones per sequencing run, sequencing was halted and the pool was resynthesized without those sequences that had already been obtained. Approximately 70 chips were reiteratively synthesized to maximize unique sequencing. Also, once 3 or more verified shRNAs were obtained for any given gene, the remaining shRNAs targeting that gene were also withdrawn from population selected for resynthesis.

To date, we have sequence verified 71,463 shRNAs targeting 28,371 human genes and 56,189 shRNAs targeting 25,831 mouse genes. A tabulation of coverage within selected functional groups can be found in Table 1 for the mouse and human libraries. The ultimate goal is to generate 3 shRNAs for each target locus. Existing, sequence-verified shRNAs for human are listed in supplementary table 1, and verified mouse shRNAs are listed in supplementary table 2. The full collection, updated dynamically, can be accessed at http:\\codex.cshl.edu.

Validation of the second-generation shRNA libraries

To test the efficiency of the second-generation shRNA libraries, we took an approach that we had previously used to assess the performance of the first generation reagents [37]. A green fluorescent protein (ZsGreen) reporter harboring the PEST domain of the mouse ornithine decarboxylase is normally degraded by the proteasome [48]. Thus, cells harboring a destabilized ZsGreen expression plasmid show very low levels of fluorescence. Interference with proteasome function, for example using a synthetic proteasome inhibitor, causes accumulation of the protein and a corresponding increase in fluorescence. The protein can also be stabilized by suppression of any gene required for proteasome function. Thus, co-transfection of the reporters with an shRNA^{mir} expression plasmid can reveal whether a target protein is involved in the proteasome pathway (Fig. 3a). Using this assay as a primary test we compared a series of shRNAs targeting proteasomal subunits that were obtained from either the first- or second-generation libraries (figure 3b).

We chose a total of 53 shRNAs targeting 13 different genes that were known to be involved in proteasome function (Fig 3b). 24 were from the first-generation library and 29 were from the second-generation library. These were co-transfected with the reporter in combination with a dsRED-encoding plasmid that allowed normalization of the transfections. It was immediately apparent that the second-generation shRNAs performed substantially better than the first-generation shRNAs. We noted that most of the plasmids derived from the second-generation library were as potent as the best shRNAs that had been selected from a screen of the first-generation library.

To gain a more detailed picture of the performance of the second-generation libraries, we compared results from the proteasome assay for 36 shRNA^{mir} expression plasmids to suppression of target RNAs as measured by semi-quantitative RT-PCR. Plasmids were transfected into HeLa cells with approximately 80% efficiency, as measured by reference to a co-transfected reporter plasmid. Despite this incomplete transfection, all but 6 of the shRNA^{mir}s reduced the levels of their target RNAs by ~60% or more with 13/36 of the shRNA^{mir}s suppressing their targets by the theoretical maximum of ~80% (Fig. 3c, upper panel). Similar results were seen with an additional 12 shRNAs that did not target proteasome subunits (not shown). These studies were also

illuminating, as they revealed that the functional assay in some cases, e.g., pSMB3, did not show a large activation of the reporter despite substantial suppression of the targeted mRNA (Fig 3c, lower panel). Thus, the functional assay underestimated slightly the efficacy of the library.

To test the performance of the library on a larger scale, we assayed a set of 515 kinase shRNAs that contained within it 47 hairpins directed to proteasome subunits using the phenotypic assay for proteasome function via a high-throughput protocol in 96-well plates (Fig. 4). In this context, 34/47 shRNAs targeting the proteasome scored as positives (72%) as compared to 10 shRNAs that had not previously been linked to proteasome function (1.9%). A secondary screen of those 44 potential positives from the primary screen again revealed positive signals from all 34 proteasomal shRNAs. However, only 5/10 of the non-proteasomal RNAs continued to activate the reporter (Supplementary table 3).

Discussion

Since the discovery that an RNAi pathway was conserved in mammals, the exploitation of this silencing response as a genetic tool has evolved in concert with our deeper understanding of its biochemical mechanism. The initial applications of siRNAs as triggers of the silencing response required comprehension of the way in which Dicer processes long dsRNA substrates in *Drosophila* [49]. Similarly, studies of dicer-mutant *C. elegans* demonstrated that endogenous loci could encode triggers of the RNAi machinery, and this led to the notion that such loci could be altered to target genes for experimental silencing [21-24]. At the time when the first such experiments were done, the nature of the primary microRNA transcript was unknown. Indeed, it was suspected that small RNAs were transcribed from the genome as short hairpin precursors, pre-miRNAs that were converted to mature small RNAs by Dicer cleavage. What has recently become clear is that pre-miRNAs are simply a processing intermediate, generated by cleavage of a longer primary transcript (pri-miRNA) by the Microprocessor [27-31].

Many strategies have been developed for producing miRNA-like triggers of the RNAi pathway. However, few studies have compared the many technical approaches to encoded RNAi triggers, directly. One impediment to such comparisons is the lack of an understanding of how each particular construction gives rise to its small RNA product. As we have mapped the processing sites on precursor shRNAs and shRNA^{mir}s, we could directly compare expression levels and efficacy for triggers carried in pSM1 to those carried in pSM2. We find that delivery of a primary microRNA expression cassette gives greater accumulation of mature small RNAs *in vivo*. Moreover, as we can predict precisely what small RNA is generated from each shRNA^{mir} expression vector, we can now apply siRNA design rules to shRNA^{mir} expression cassettes. A combination of increased small RNA production with better shRNA design yielded a pronounced increase in the performance of these silencing tools.

Guided by these design strategies, we have constructed large libraries of sequence-verified shRNAs targeting 2/3 of the known and predicted genes in the human and mouse genomes. On average, each locus is covered by 2 shRNA^{mir}s presently; however, the ultimate goal is to have 3 sequence-verified shRNA^{mir}s for each gene. The second generation libraries resemble those that we have previously reported in that they reside in flexible vectors that permit shuttling of shRNA^{mir} expression cassettes into virtually any desired expression vector using a bacterial mating strategy [46]. A unique feature of the second generation library is that the expression cassette can be moved without the need to move also the constitutive U6 snRNA promoter. This permits large scale construction of secondary libraries under the control of tissue specific and inducible promoters. Indeed, regulated expression of our library cassettes from RNA polymerase II promoters has been shown to effectively suppress gene expression both in cultured cells and in animals (Dickins et al, see accompanying paper; Stegmeier et al., submitted). These recipient vectors can be directly used with any shRNA^{mir} encoded by the library described herein.

The use of large-scale resources for suppressing gene expression via RNAi promises to revolutionize genetic approaches to biological problems in numerous model systems. For human cells, both siRNA and shRNA collections have previously been reported and are, to varying degrees, generally available to investigators to probe a wide range of biological questions. The libraries described here should prove useful for assessing the functions of individual genes and for taking genome-wide approaches. Strategies reported in the accompanying paper and by Stegmeier and colleagues will permit large-scale application of these tools for screens which require long-term suppression of gene expression using single-copy integrants or inducible repression. Thus, we have produced coherent system of RNAi reagents with utility in both mouse and human experimental systems.

Methods

Construction of the lysogenic strain DH10βλ_{KP} and excision strain BUN25
DH10βλ_{KP} [*mcrA Δ(mrr-hsdRMS-mcrBC) ϕ80 lacZΔM15 ΔlacX74 deoR recA1 endA1 araΔ139 Δ(ara, leu)7697 galU galK λ⁻ rpsL nupG tonA λ-pir1-npt*] is a strain containing λcI and the *pir1* gene that was constructed in order to lysogenize the λSM2-barcode library prior to introduction of the hairpin fragments. To generate this strain, λ_{KP} containing the *pir1* and Kn^R genes was constructed. To generate λ_{KP}, the *pir1* gene was amplified from BW23473 using primers MZL393 and MZL51, and cloned into the pCR2.1 TOPO TA cloning vector. The *pir1* gene was excised from the above clone on a BamH1 fragment and ligated into BamH1 cleaved pSE356, which contains an *npt* gene and a *Bam*HI restriction site flanked by two 1 kb λ DNA fragments [50] to generate pSE356pirWT. The pir1Km^R fragment was recombined onto wild type λ by amplifying λon LE392/pSE356pirWT and the resulting phage were collected and used to infect DH10β. 100 μl of DH10β cells were infected with 10^6 PFU at 30°C for 30 minutes in LB + 10 mM $MgSO_4$, diluted with 900 μl of LB incubated at 30°C for 2 h with shaking, and plated on LB containing 50 μg/ml kanamycin at 37°C overnight to select λ_{KP} lysogens. Lysogens were tested for the ability to lysogenize λ vectors containing R6Kγ origins of replication as extrachromosomal elements. A strain capable of doing this was selected and named DH10βλ_{KP}.

The BUN25 [F' *traD36 lacI^q Δ(lacZ)M15 proA$^+$B$^+$/e14$^-$(McrA$^-$) Δ(lac-proAB) thi gyrA96* (Nal^r) *endA1 hsdR17 ($r_k^-m_k^+$) relA1 glnV44 λ-cre-npt umuC::pir116-Frt sbcDC-Frt*] strain containing *pir1-116* and *cre* was constructed to allow the conversion of λSM2 shRNA libraries into pSM2 shRNA libraries. A PCR fragment containing *pir1-116* gene was generated using primers MZL393 and MZL51 (see Supplementary table 4), cleaved with BamH1 and ligated into *Bam*HI-cleaved pUC18 to generate pML284. A fragment containing *Bst*BI-Frt-*cat*-Frt-*Nde*I (filled-in) was isolated from KD3 [51] and inserted into the *Sma*I site of pML284 to generate pML334. A *Hpa*I fragment containing *UmuDC* was isolated from pSE117 and cloned into pBluescript *Xho*I (filled in)-*Eco*RV to generate pML236. We eliminated one of the *Bam*HI site on pML236 by digesting it with *Pst*I-*Xba*I, filling in with T4 DNA polymerase and ligating. The Frt-*cat*-Frt-*pir116* was isolated from pML334 as a *Kpn*I-*Sac*I (filled-in) fragment and ligated into *Mlu*I/*Bam*HI (filled-in) cleaved pML236-ΔBamHI to generate pML346. The 3.8 kb *Kpn*I-*Sac*I *UmuDC*-Frt-*cat*-Frt-*pir116-UmuC* fragment from pML346 was integrated into BNN132/pML104 by homologous recombination using the λ recombinase expressed from pML104, and confirmed by colony PCR. The *cat* gene was removed by FLP-mediated excision *in vivo* using pCP20 [52] which expresses the FLP recombinase to generate BUN24. A cassette that has Frt-*cat*-Frt flanked by 50 bp homology to *sbcD* and 50 bp homology to *sbcC* was amplified by primers MZL493/MZL494 and using KD3 as a template. This cassette was used to replace *sbcD* and part of *sbcC* on BNN132 by homologous recombination and the deletion were confirmed by colony PCR. The strain was

named BNN132sbcDC-Frt-cat-Frt. We then used a pair of outside primers (MZL495/MZL496) that gave about 500 bp homology regions to the upstream of *sbcD* and 500 bp homology regions to the *sbcC* to amplify a PCR product form BNN132sbcDC-Frt-cat-Frt to recombine onto the *sbcDC* region of BUN24. The resulting strain was named BUN25 and is used to stabilize inverted repeats in E. coli [53].

Library vector construction

A pair of loxP-*Not*I-loxP duplexed oligos (MZL524/MZL525) were inserted into the pSM2 *Bst*XI site to generate pSM2c-loxP. A second pair of duplexed oligos (MZL541/ MZL542), carrying the proper restriction sites for cloning barcodes into λSM2, were inserted into the *Bbs*I-*Mlu*I sites of pSM2c-loxP to create pML375. λACT2 was digested with *Not*I, and the λ arms were gel purified and ligated to *Not*I digested pML375 to generate λSM2. The ligation mixture was packaged using MaxPlax™ lambda packaging extracts from Epicentre. We selected a λSM2 lysogen by infecting 200 μl of BW23473 cell (A_{600} = ~0.8) with 100 μl of λSM2 packaging mix in the presence of 10 mM $MgSO_4$ and 0.2 % (w/v) maltose, incubated at 30°C for 30 minutes the added 900 μl of LB and incubated at 30°C for 2 hs with shaking to express the Cm^R marker, and plated on LB containing 17 μg/ml of chloramphenicol (Cm) at 30°C overnight. The proper recombinants were confirmed by restriction analysis.

Barcode library construction

The 60 base pair barcodes, gaagactaatgcggccggcca$(n)_{60}$gggccctatagtgagtcgtatta, were amplified using barcode primer 1 (aaattgcaatgaagactaatgcggccggcca) and barcode primer 2 (atatatggacgcgtcctaggtaatacgactcactatagggccc). The PCR conditions were: 0.1 pmol of barcodes, 50 pmol of each primer, 25 nmol of each dNTP, and 2.5 U of Taq DNA polymerase; 94°C for 45 seconds, (94°C for 30 seconds, 55°C for 30 seconds, and 72°C for 30 seconds) x 13, 72°C for 10 minutes, 4°C forever. Ten PCR reactions were pooled together, purified using a QIAquick PCR purification kit, digested with *Bam*HI, *Eco*RI, *Xho*I and *Sal*I to remove these sites in the barcodes, and gel purified. The purified barcodes were digested with *Fse*I and *Avr*II and purified using the QIAquick gel extraction kit. Two micrograms of *Fse*I-*Avr*II digested λSM2 ligated with 10 ng of *Fse*I-*Avr*II digested barcodes with 1 x ligation buffer and 0.5 μl T4 DNA ligase in a 5 μl final volume at 16°C overnight. The ligation mixture was packaged and amplified. The size of the λSM2-barcode library was 4.2 x 10^7. We used 20 ml of DH10βλ_{KP} cells (A_{600} = ~1) to lysogenize 2 x 10^9 of λSM2-barcode library as 42 kb plasmids. The cell and the phage were mixed in the presence of 10 mM $MgSO_4$ and 0.2 % (w/v) maltose and incubated at 30°C for 30 minutes, added 200 ml of LB to recover at 30°C for 1 h by shaking. The mixture were concentrated by centrifugation at 4000 rpm for 20 minutes, resuspended in 3 ml of LB, plated on 10 large LB/Cm 17 μg/ml, and incubated at 30°C overnight. The cells were scraped from plates and grown in 3 L of TB containing 17 μg/ml of Cm overnight. Supercoiled λSM2-barcode library DNA

was prepared by cesium chloride. The lysogenization efficiency was approximately 30%.

Oligonucleotide Cleavage and PCR Amplification

To harvest oligonucleotides, we treated microarrays for 2 h with 2-3 mL of 35% NH_4OH solution (Fisher Scientific) at room temperature. We transferred the solution to 1.5-mL microcentrifuge tubes and subjected it to speed vacuum drying at medium heat (~55 °C) overnight. We resuspended the dried material in 200 µl of RNase/DNase free water and performed PCR amplification in 50µl reaction volumes using Invitrogen's Platinum® *Pfx* DNA Polymerase. To obtain a sufficient amount of PCR product, four 50 µl reactions were required for each sample. Each reaction contained 2X *Pfx* PCR amplification buffer, 0.3 mM of each dNTP, 1 mM $MgSO_4$, 0.3 uM of each primer, 0.5X PCR enhancer solution, 0.5 units of Platinum® *Pfx* DNA Polymerase, and 10 µl of template DNA. The primers used for amplification were 5'-mir30-PCR-xhoI-F (5' CAGAGGCTCGAGAAGGTATATTGCTGTTGACAGTGAGCG 3') and 3'-mir30-PCR-ecorI-R (5' CGCGGCGAATTCCGAGGAGTAGGCA 3'). After an initial denaturation step of 94°C for 5 min, DNA amplification occurred through 25 cycles of denaturing at 94°C for 45 seconds and annealing and extension at 68° for 1 min and 15 sec followed by a final 7 minute extension at 68°. We then combined the four reactions into one tube, cleaned up the PCR product by use the QIAGEN® Minelute PCR Purification Kit and eluted in a total volume of 26 µl.

ShRNA library construction

The λSM2-barcode library and shRNA PCR products were digested with *Eco*RI and *Xho*I overnight and gel purified. Ligations with shDNA oligos were set up as following: 1.5 µg of *Xho*I-*Eco*RI cleaved vector, 8-10 ng of *Xho*I-*Eco*RI cleaved inserts generated from the PCR of shDNA oligos from the parallel microarray synthesis, 1 µl of 10 x ligation buffer, 0.5 µl of T4 DNA ligase, and water to 10 µl final. The ligation mixtures were incubated at 16°C for overnight and packaged. We typically observed 30- to 90-fold stimulation of plaque forming units (PFU) and 2 x 10^7 to 8 x 10^7 PFU total for each library pool. We typically amplify 2 x 10^7 PFU for each pool to generate a stock. To verify the ligation efficiency, we excised 10 µl of package mix by infecting 100 µl of BUN25 (A = ~0.5) and selected colonies on LB/Cm 30 µg/ml 30°C overnight. Colony PCR was performed using forward (ggacgaaacaccgtgctcgc) and reverse primer (ttctgcgaagtgatcttccg) and 85 to 95 % correct sized inserts were typically observed with some containing multiple inserts. To generate plasmid DNA from these libraries, we typically excised 5 x 10^7 PFUs through infection of BUN25 cells as described earlier. The cells were scraped from plates and grown in 2 L of LB plus 13 g/l of circle growth 37°C for 7 to 8 h. Cesium chloride method was used to prepare DNA. DNAs were transformed into BW23474 F'DOT SbcC , and individual clones were sequenced using primer5' (TGTGGAAAGGACGAAACACC). Correct clones were individually rearrayed to form the final library.

Small RNA Northern Blots
293 cells were transfected in 10cm dishes at 60% confluency with 15 ug of shRNA plasmid DNA along with 5 ug of pDsRed-N1 (Clontech) using TransIT-LT1 (Mirus). 48hs post transfection, transfection efficiency was confirmed by estimating the percentage of cells expressing DsRed (~80%) and then total RNA was Trizol extracted and purified. Small RNA northern blots were carried out as described in [54] using 30 ug RNA/lane. For hairpins targeting EGFP at starting position 481 (Fig 1b), northern probes were DNA oligos corresponding to the anti-sense strand (ccggcatcaaggtgaacttcaa) of the mature RNA.

Proteasome assays
Bacteria cultures were grown in 96 well plates for 36 h in GS96 media (Bio101). Plasmid DNA was extracted using Quiagen Ultrapure plasmids minipreps in a 96 well plate format. DNA concentrations were determined by mixing an aliquot of each sample with picogreen (Molecular Probes) and determining fluorescence on a Victor2 plate reader. HEK 293T cells were plated in 96 well optic plates (Corning) at $1x10^5$ cells per ml. For the proteasome assay, 12.5 ng of the plasmid dsRed N-1 (Clontech), 5 ng of the Zsprosensor (Clontech) and 75 ng of each individual shRNA construct were cotransfected per well using 0.3 μl of LT-1 (Mirus) transfection reagent. After 24 hs the transfection media was replaced. After 72 h, media was removed and replaced with PBS in order to read fluorescence. Fluorescence signals were read on a Victor2 plate reader. Signals in the green channel were normalized to transfection efficiency using customized scripts with fluorescence in the red channel serving as a normalization criterion. Cut-offs were assigned by using control shRNA transfections to determine the range for a negative outcome.

Plasmid Transfections and mRNA Quantitation
HeLa cells were seeded at 0.5×10^5 cells/well in 24-well plates and transfected 24 h later with 1 ug/well of the appropriate plasmid. Each plasmid was delivered to 4 wells by use of Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocols. Transfection efficiency was determined by parallel transfection of a GFP-expressing plasmid and the percentage of fluorescent cells assayed by flow cytometry. For analysis of target gene mRNA knock down, cell lysates were collected 24 h after transfection, and total RNA was prepared by use of RNeasy columns (Qiagen) following the manufacturer's protocols. Messenger RNA quantitation was performed by Real-time PCR of reverse transcription products, using available Applied Biosystems TaqMan™ primer probe sets, and the percent mRNA remaining was determined by comparison with mRNA levels from cells treated with transfection reagent alone.

Acknowledgements

We thank members of the Hannon, Elledge, and Lowe labs for helpful suggestions. Jill Magnus assisted with array PCRs, and Lidia Nascimento, Vivek Balija, Melissa Kramer, Theresa Zutavern, Stephanie Muller, and Beth Miller

assisted with sequencing. Troy Moore helped with curation of the collection. We thank Peter Linsley and Steve Friend for their support of this project. This work was funded in part by awards from the DOD Breast Cancer Research Program (GJH and SJE) and the NIH (GJH and SJE).

Table 1.

Functional Groups	Human Genes in the category	Human shRNA clones	Mouse Genes in the category	Mouse shRNA clones
Apoptosis	576	1948	537	1395
Cancer relevant	846	2888	761	2043
Cell Cycle	519	2055	461	1324
Checkpoint	124	526	111	320
DNA Repair	118	509	104	275
DNA Replication	234	922	223	592
Enzymes	2876	9846	2550	6871
GPCR	652	1950	574	1436
Kinases	625	2614	566	2149
Phosphatases	199	712	173	536
Proteases	444	1340	401	988
Proteolysis	310	1427	237	681
Signal Transduction	2610	8541	2319	6101
Trafficking	473	1487	403	1013
Transcription	832	2734	676	1808

Figure Legends

Figure 1. Design and structure of shRNA^{mir} cassettes. (**a**) A comparison of the structures of several silencing triggers is shown. These include an siRNA, a portion of the shRNA precursor, as generated from our first-generation design in pSM1, and a segment of the shRNA^{mir} precursor produced by pSM2. The sequence of the target site (sense orientation) from firefly luciferase (luc1309, see **c**) is shown in red. For pSM2, cleavage sites for Dicer and Drosha, as mapped by primer extension, are indicated by arrows. (**b**) Northern blotting was used to detect the mature small RNA produced after transfection of HEK-293T cells with shRNA and shRNA^{mir} cassettes expressed from pSM1 and pSM2, resptectively, by the U6 snRNA promoter. In neither case was significant accumulation of pre-miRNA observed. Transfection rates were normalized using a co-delivered dsRED expression plasmid. (**c**) Five different promoters (human tRNAval, Human H1 RNA, Human U6 snRNA, MSCV LTR and Human CMV IE, as indicated) were tested for their ability to drive shRNA^{mir} expression and silence luciferase in transient transfections. Two different shRNAs were used, a highly efficient shRNA (luc1309) an a less efficient shRNA (luc311). In each case, the level of firefly luciferase was normalized to a non-targeted *Renilla* luciferase.

Figure 2. Construction of the second-generation library. (a) The pSM2c vector contains a U6 promoter, a U6 terminator following mir3', a self inactivating retroviral backbone; two bacterial antibiotic resistance markers kanamycin and chloramphenicol; a protein-dependent origin (RK6γ); a mammalian selectable marker (puromycin) driven by the PGK promoter; a homology region (HR2) for use in bacterial recombination and a randomly generated 60 mer barcode sequence. The shRNA^{mir} inserts were cloned between the 5' and 3' flanking sequences derived from the mir-30 primary transcript using XhoI and EcoRI restriction sites. The nucleotide positions for sites in an excised version of an empty vector (no shRNA or barcode) are given. **(b)** Construction of the second-generation libraries began with the generation of a lambda derivative of pSM2 that contained unique EcoR1, XhoI, FseI and AvrII sites, the latter two for insertion of bar codes. A bar coded pre-library was generated by the ligation of PCR amplified random 60 mers into Fse1-AvrII cleaved λpSM2 to generate a bar-coded library pool (upper right). The bar-codes λpSM2 was converted into a shRNA library by insertion of PCR amplified shRNA constructs prepared by *in situ* synthesis of inserts on a microarray in pools of 22,000 into the EcoR1-XhoI cleaved pre-library (upper right). Packaged phage were amplified and used to infect BUN25, which express Cre recombinase and *pir1-116* for pSM2 replication. Each excision event gave rise to a Kan^r+Cm^r resistant colony. These were pooled and used for preparation of library DNA. This, in turn, was transformed into BW F'DOT and individual colonies were selected for sequence analysis.

Figure 3. Validation of the second-generation library. (a) A schematic representation of the phenotypic assay for proteasome function (see text) is

shown. **(b)** Thirteen proteasome subunits were chosen because of their representation in both the first- and second-generation libraries. ShRNA expression clones corresponding to each were assayed for activation of the proteasome reporter. Blue bars indicate first-generation clones while green indicate second-generation clones. In all cases, the activity of the proteasome reporter (green channel) was normalized for transfection using a dsRED expression plasmid. **(c)** In a separate study, 36 different proteasome shRNAs were tested for their ability to suppress their target RNAs (upper panel). QRT-PCRs were performed 24 h after transfection of HeLa cells at an average efficiency of 80% as measured by a co-transfected normalization reporter (dsRED). For comparison, functional assays for proteasome inhibition were performed in parallel (lower panel).

Figure 4. Performance of the second-generation library in a small-scale high-throughput screen. 47 shRNAs targeting proteasome subunits were distributed among a series of 562 hairpins targeting human kinases (upper panel). The lower left panel shows the negative (FF) and the positives controls (ATPase 1.1 to 1.3) from first (pSM1) and second (pSM2) libraries. The lower right panel shows the shRNAs that displayed accumulation of the proteasome reporter over the cut-off (2-fold or greater activation; yellow line). These are highly enriched for proteasome shRNAs (red). In blue are 10 additional non-proteasomal shRNAs that also scored positive in the screen. Of these, 5 were also positive on a retest of individual clones (see Supplementary Table 3).

Literature Cited

1. Hannon, G.J. RNA interference. *Nature* **418**, 244-51. (2002).
2. Tomari, Y. & Zamore, P.D. Perspective: machines for RNAi. *Genes Dev* **19**, 517-29 (2005).
3. Nakayashiki, H. et al. RNA silencing as a tool for exploring gene function in ascomycete fungi. *Fungal Genet Biol* **42**, 275-83 (2005).
4. Tang, G. & Galili, G. Using RNAi to improve plant nutritional value: from mechanism to application. *Trends Biotechnol* **22**, 463-9 (2004).
5. Dasgupta, R. & Perrimon, N. Using RNAi to catch Drosophila genes in a web of interactions: insights into cancer research. *Oncogene* **23**, 8359-65 (2004).
6. Fraser, A. Towards full employment: using RNAi to find roles for the redundant. *Oncogene* **23**, 8346-52 (2004).
7. Silva, J., Chang, K., Hannon, G.J. & Rivas, F.V. RNA-interference-based functional genomics in mammalian cells: reverse genetics coming of age. *Oncogene* **23**, 8401-9 (2004).
8. Elbashir, S.M. et al. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494-8. (2001).
9. Donze, O. & Picard, D. RNA interference in mammalian cells using siRNAs synthesized with T7 RNA polymerase. *Nucleic Acids Res* **30**, e46 (2002).
10. Kawasaki, H., Suyama, E., Iyo, M. & Taira, K. siRNAs generated by recombinant human Dicer induce specific and significant but target site-independent gene silencing in human cells. *Nucleic Acids Res* **31**, 981-7 (2003).
11. Kittler, R. et al. An endoribonuclease-prepared siRNA screen in human cells identifies genes essential for cell division. *Nature* **432**, 1036-40 (2004).
12. Luo, B., Heard, A.D. & Lodish, H.F. Small interfering RNA production by enzymatic engineering of DNA (SPEED). *Proc Natl Acad Sci U S A* **101**, 5494-9 (2004).
13. Hammond, S.M., Boettcher, S., Caudy, A.A., Kobayashi, R. & Hannon, G.J. Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science* **293**, 1146-50. (2001).
14. Song, J.J. et al. The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nat Struct Biol* **10**, 1026-32 (2003).
15. Song, J.J., Smith, S.K., Hannon, G.J. & Joshua-Tor, L. Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* **305**, 1434-7 (2004).
16. Liu, J. et al. Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305**, 1437-41 (2004).
17. Parker, J.S., Roe, S.M. & Barford, D. Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition and slicer activity. *Embo J* **23**, 4727-37 (2004).

18. Bartel, D.P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281-97 (2004).
19. He, L. & Hannon, G.J. MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet* **5**, 522-31 (2004).
20. Reinhart, B.J. et al. The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature* **403**, 901-6. (2000).
21. Ketting, R.F. et al. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. *Genes Dev* **15**, 2654-9. (2001).
22. Grishok, A. et al. Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. *Cell* **106**, 23-34. (2001).
23. Knight, S.W. & Bass, B.L. A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in Caenorhabditis elegans. *Science* **293**, 2269-71. (2001).
24. Hutvagner, G. et al. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* **293**, 834-8. (2001).
25. Lee, Y. et al. MicroRNA genes are transcribed by RNA polymerase II. *Embo J* **23**, 4051-60 (2004).
26. Cai, X., Hagedorn, C.H. & Cullen, B.R. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *Rna* **10**, 1957-66 (2004).
27. Lee, Y. et al. The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415-9 (2003).
28. Denli, A.M., Tops, B.B., Plasterk, R.H., Ketting, R.F. & Hannon, G.J. Processing of primary microRNAs by the Microprocessor complex. *Nature* **432**, 231-5 (2004).
29. Landthaler, M., Yalcin, A. & Tuschl, T. The Human DiGeorge Syndrome Critical Region Gene 8 and Its D. melanogaster Homolog Are Required for miRNA Biogenesis. *Curr Biol* **14**, 2162-7 (2004).
30. Han, J. et al. The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev* **18**, 3016-27 (2004).
31. Gregory, R.I. et al. The Microprocessor complex mediates the genesis of microRNAs. *Nature* **432**, 235-40 (2004).
32. Yi, R., Qin, Y., Macara, I.G. & Cullen, B.R. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev* **17**, 3011-6 (2003).
33. Lund, E., Guttinger, S., Calado, A., Dahlberg, J.E. & Kutay, U. Nuclear export of microRNA precursors. *Science* **303**, 95-8 (2004).
34. Siolas, D. et al. Synthetic shRNAs as potent RNAi triggers. *Nat Biotechnol* **23**, 227-31 (2005).
35. Schwarz, D.S. et al. Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **115**, 199-208 (2003).
36. Khvorova, A., Reynolds, A. & Jayasena, S.D. Functional siRNAs and miRNAs exhibit strand bias. *Cell* **115**, 209-16 (2003).

37. Paddison, P.J. et al. A resource for large-scale RNA-interference-based screens in mammals. *Nature* **428**, 427-31 (2004).
38. Berns, K. et al. A large-scale RNAi screen in human cells identifies new components of the p53 pathway. *Nature* **428**, 431-7 (2004).
39. Zeng, Y., Wagner, E.J. & Cullen, B.R. Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. *Mol Cell* **9**, 1327-33. (2002).
40. Parizotto, E.A., Dunoyer, P., Rahm, N., Himber, C. & Voinnet, O. In vivo investigation of the transcription, processing, endonucleolytic activity, and functional relevance of the spatial distribution of a plant miRNA. *Genes Dev* **18**, 2237-42 (2004).
41. Paddison, P.J., Caudy, A.A., Bernstein, E., Hannon, G.J. & Conklin, D.S. Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev* **16**, 948-58. (2002).
42. Chen, C.Z., Li, L., Lodish, H.F. & Bartel, D.P. MicroRNAs modulate hematopoietic lineage differentiation. *Science* **303**, 83-6 (2004).
43. Kawasaki, H. & Taira, K. Short hairpin type of dsRNAs that are controlled by tRNA(Val) promoter significantly induce RNAi-mediated gene silencing in the cytoplasm of human cells. *Nucleic Acids Res* **31**, 700-7. (2003).
44. Brummelkamp, T.R., Bernards, R. & Agami, R. A System for Stable Expression of Short Interfering RNAs in Mammalian Cells. *Science* **21**, 21 (2002).
45. Zheng, L. et al. An approach to genomewide screens of expressed small interfering RNAs in mammalian cells. *Proc Natl Acad Sci U S A* **101**, 135-40 (2004).
46. Li, M.Z. & Elledge, S.J. MAGIC, an in vivo genetic method for the rapid construction of recombinant DNA molecules. *Nat Genet* **37**, 311-9 (2005).
47. Cleary, M.A. et al. Production of complex nucleic acid libraries using highly parallel in situ oligonucleotide synthesis. *Nat Methods* **1**, 241-8 (2004).
48. Li, X. et al. Generation of destabilized green fluorescent protein as a transcription reporter. *J Biol Chem* **273**, 34970-5 (1998).
49. Carmell, M.A. & Hannon, G.J. RNase III enzymes and the initiation of gene silencing. *Nat Struct Mol Biol* **11**, 214-8 (2004).
50. Elledge, S.J. & Walker, G.C. Phasmid vectors for identification of genes by complementation of Escherichia coli mutants. *J Bacteriol* **162**, 777-83 (1985).
51. Datsenko, K.A. & Wanner, B.L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**, 6640-5 (2000).
52. Cherepanov, P.P. & Wackernagel, W. Gene disruption in Escherichia coli: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene* **158**, 9-14 (1995).
53. Chalker, A.F., Leach, D.R. & Lloyd, R.G. Escherichia coli sbcC mutants permit stable propagation of DNA replicons containing a long palindrome. *Gene* **71**, 201-5 (1988).

54. Caudy, A.A., Myers, M., Hannon, G.J. & Hammond, S.M. Fragile X-related protein and VIG associate with the RNA interference machinery. *Genes Dev* **16**, 2491-6. (2002).

Suplementary material

Supplementary Table 1. Content of the human shRNA library

Supplementary Table 2. Content of the mouse shRNA library

Available on accompanying CD

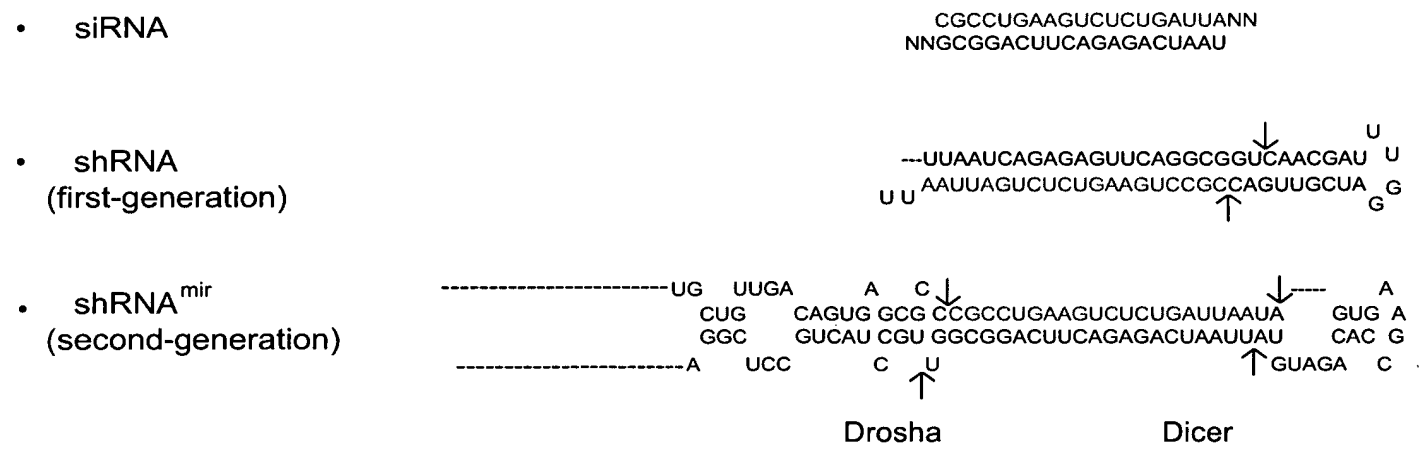
Supplementary table 3. Non-proteasomal genes identified as positives

Acc.Number	Name
XM_114099	Homo sapiens similar to adenylate kinase 1 (LOC200047)
NM_004570	Homo sapiens phosphoinositide-3-kinase, class 2, gamma polypeptide
NM_007284	Homo sapiens PTK9L protein tyrosine kinase 9-like
NM_002576	Homo sapiens p21/Cdc42/Rac1-activated kinase 1
NM_006622	Homo sapiens polo-like kinase 2 (PLK2)

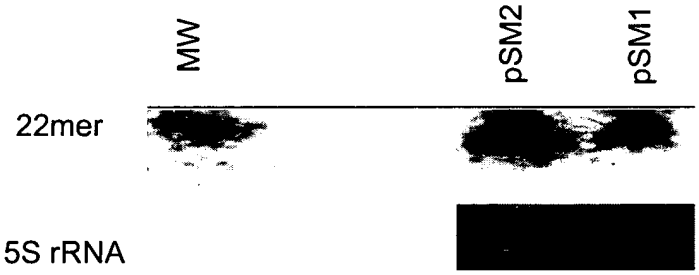
Supplementary table 4. Oligonucleotides used in construction of the library vectors

oligo names	oligo sequences
MZL393	AAATTTGGATCCGCGTTAAACATGAGTGGATAG
MZL51	TATATAGGATCCGGCGTTCGGCTCCTTGAGGG
MZL493	GCAAATGGCACATCTGTTTGGGTATAATCGCGCCCATGCTTTTTCGCCA GGTGTAGGCTGGAGCTGCTTCG
MZL494	CCGCGCTGTGTTCTGCGATGCGTTCCCAGTGTGGACGAAGATTTCGTGC CGGCATATGAATATCCTCCTTAGTT
MZL495	AATTAAGGATCCGCACGCGATCTTCTTCTTCC
MZL496	ATATATAAGCTTTGACATTCTGGATAGCGACC
MZL524	ATAACTTCGTATAGCATACATTATACGAAGTTATGCGGCCGCATAA CTTCGTATAGCATACATTATACGAAGTTATGCTG
MZL525	ATAACTTCGTATAATGTATGCTATACGAAGTTATGCGGCCGCATAACTTC GTATAATGTATGCTATACGAAGTTATCAGC
MZL541	ATGCGGCCGGCCGTTTAAACATTTAAATCCTAGGA
MZL542	CGCGTCCTAGGATTTAAATGTTTAAACGGCCGGCC
Barcodes	GAAGACTAATGCGGCCGGCCA $(N)_{60}$GGGCCCTATAGTGAGTCGTATTA
Barcode primer 1	AAATTGCAATGAAGACTAATGCGGCCGGCCA
Barcode primer 2	ATATATGGACGCGTCCTAGGTAATACGACTCACTATAGGGCCC
Forward primer (for colony PCR)	GGACGAAACACCGTGCTCGC
Reverse primer (for colony PCR)	TTCTGCGAAGTGATCTTCCG

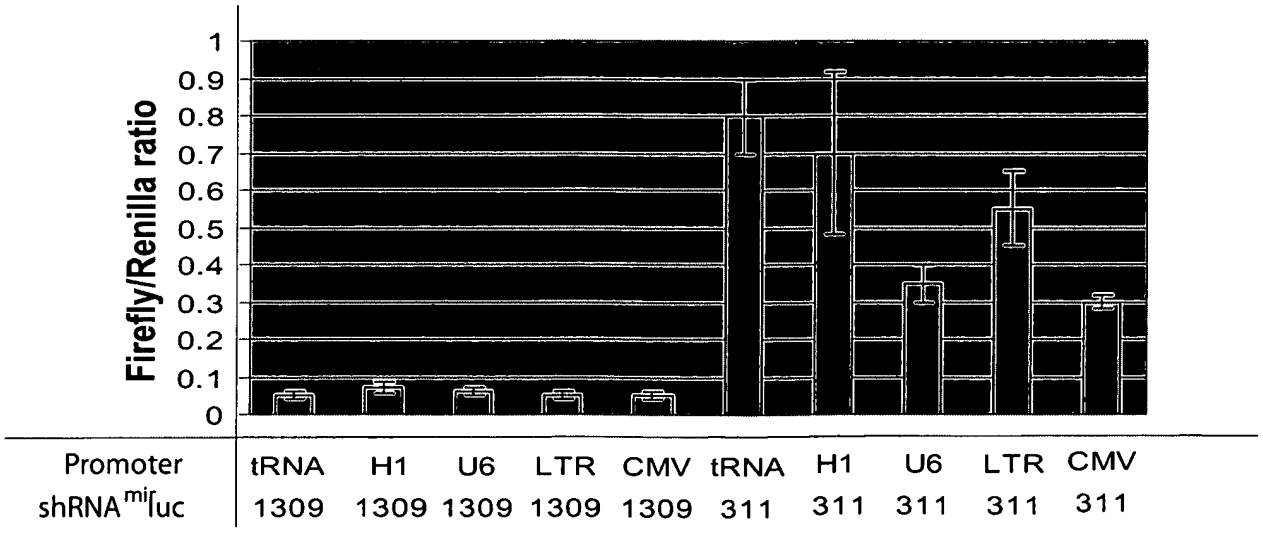
a


siRNA
CGCCUGAAGUCUCUGAUUANN
NNGCGGACUUCAGAGACUAAU
shRNA
(first-generation)
---UUAAUCAGAGAGUUCAGGCGGUCAACGAU U
U U AAUUAGUCUCUGAAGUCCGCCAGUUGCUA G
G
shRNA mir
(second-generation)
----------UG UUGA A C
CUG CAGUG GCG CCGCCUGAAGUCUCUGAUUAAUA GUG A
GGC GUCAU CGU GGCGGACUUCAGAGACUAAUUAU CAC G
----------A UCC C U
GUAGA C
Drosha
Dicer

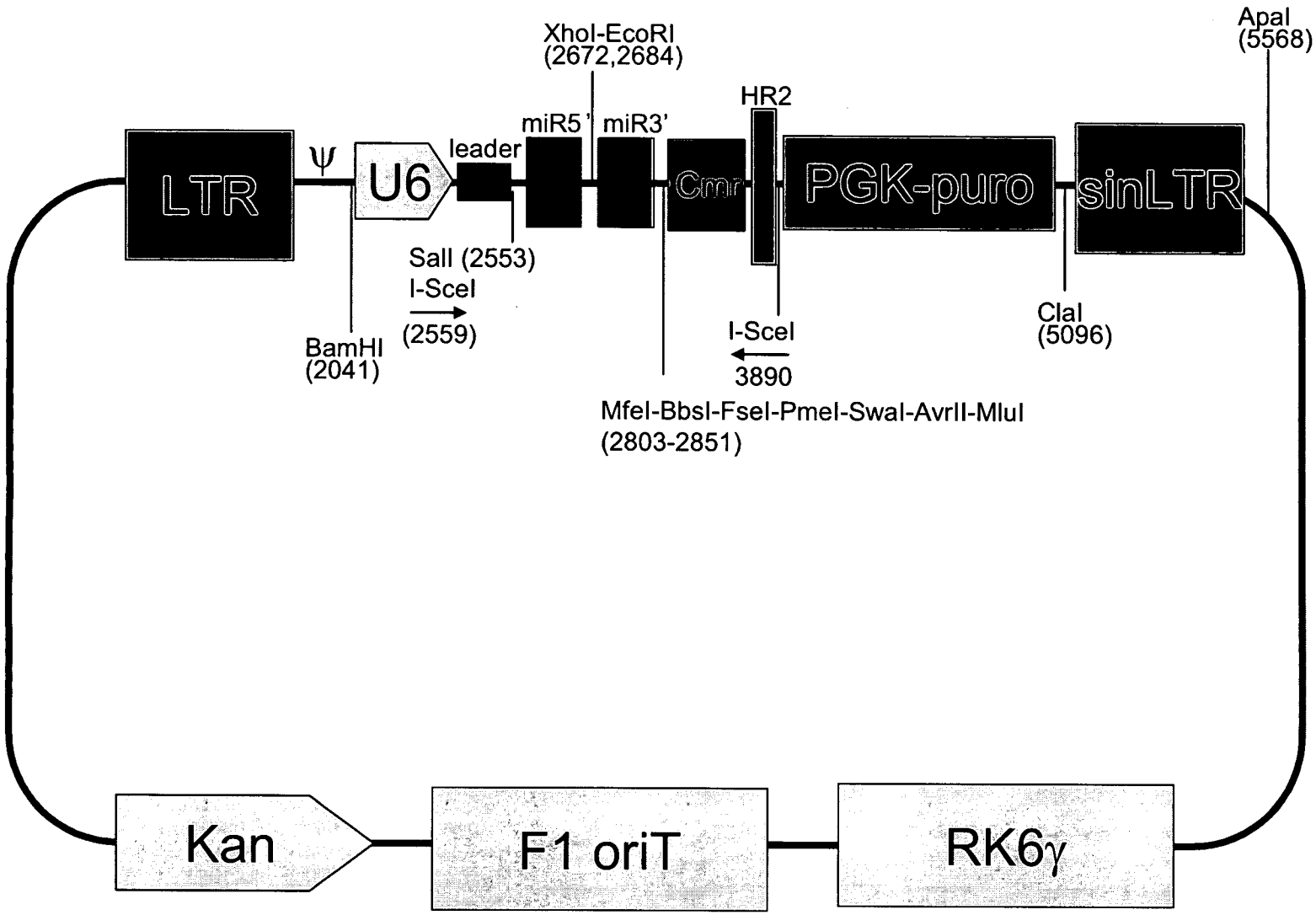
b

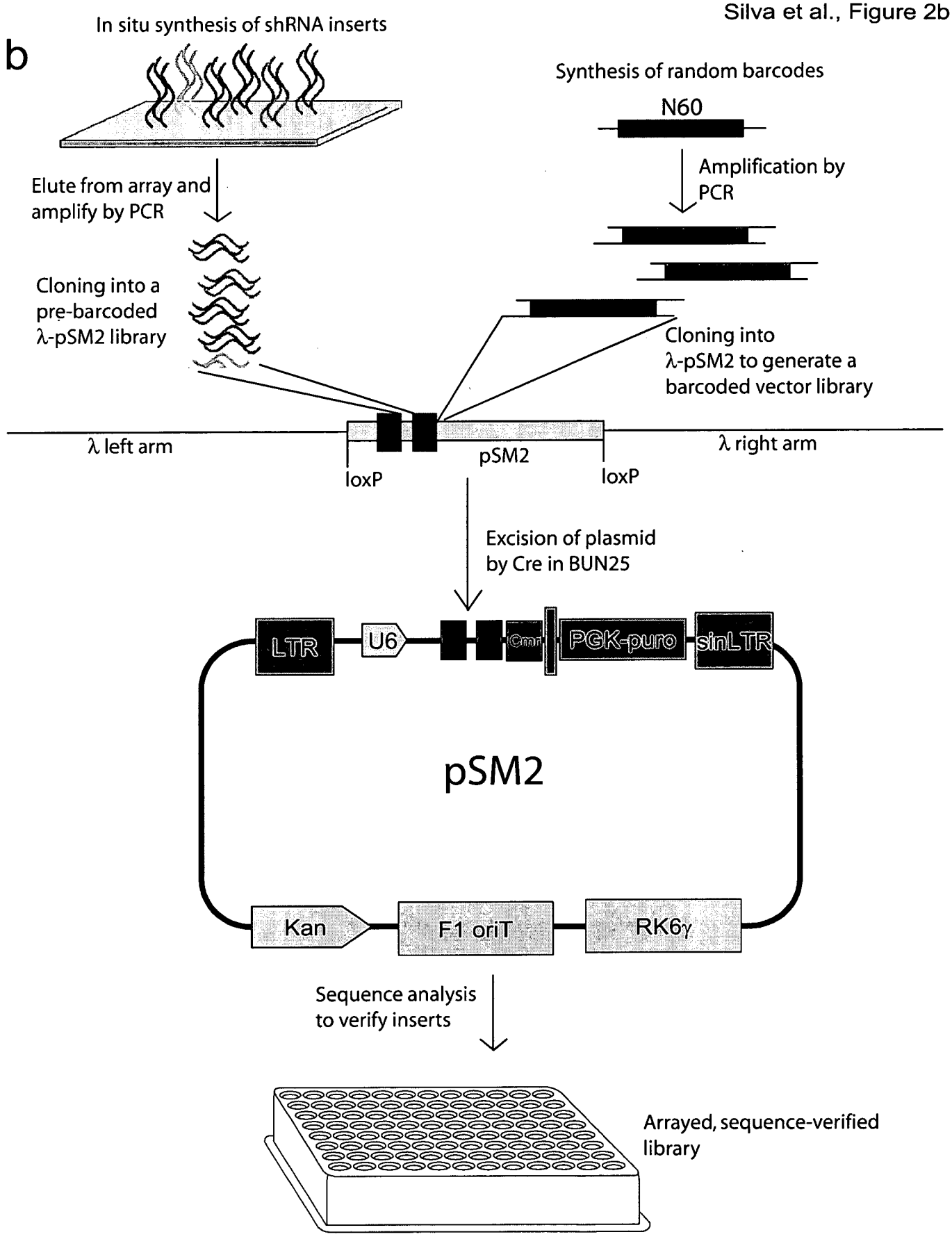

MW
pSM2
pSM1
22mer
5S rRNA

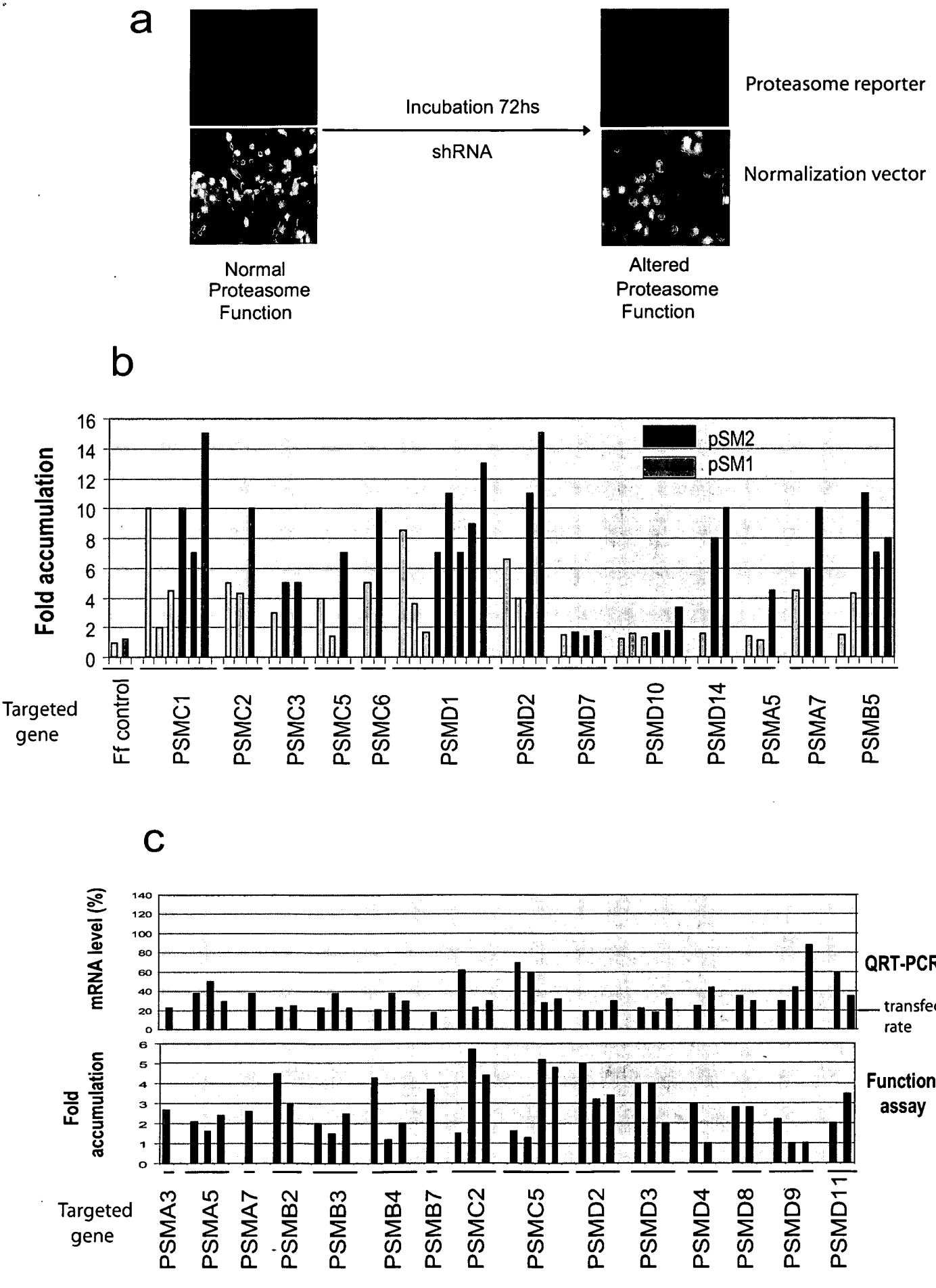
c


Firefly/Renilla ratio
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Promoter tRNA H1 U6 LTR CMV tRNA H1 U6 LTR CMV
shRNA mir luc 1309 1309 1309 1309 1309 311 311 311 311 311

a


XhoI-EcoRI
(2672,2684)
ApaI
(5568)
HR2
miR5'
miR3'
leader
LTR
Ψ
U6
Cmr
PGK-puro
sinLTR
SalI (2553)
I-SceI
(2559)
BamHI
(2041)
I-SceI
3890
ClaI
(5096)
MfeI-BbsI-FseI-PmeI-SwaI-AvrII-MluI
(2803-2851)
Kan
F1 oriT
RK6γ


b
In situ synthesis of shRNA inserts
Synthesis of random barcodes
N60
Amplification by PCR
Elute from array and amplify by PCR
Cloning into a pre-barcoded λ-pSM2 library
Cloning into λ-pSM2 to generate a barcoded vector library
λ left arm
pSM2
λ right arm
loxP
loxP
Excision of plasmid by Cre in BUN25
LTR
U6
Cmr
PGK-puro
sinLTR
pSM2
Kan
F1 oriT
RK6γ
Sequence analysis to verify inserts
Arrayed, sequence-verified library


a
Incubation 72hs
shRNA
Proteasome reporter
Normalization vector
Normal Proteasome Function
Altered Proteasome Function
b
Fold accumulation
pSM2
pSM1
Targeted gene
Ff control
PSMC1
PSMC2
PSMC3
PSMC5
PSMC6
PSMD1
PSMD2
PSMD7
PSMD10
PSMD14
PSMA5
PSMA7
PSMB5
c
mRNA level (%)
QRT-PCR
transfection rate
Fold accumulation
Functional assay
Targeted gene
PSMA3
PSMA5
PSMA7
PSMB2
PSMB3
PSMB4
PSMB7
PSMC2
PSMC5
PSMD2
PSMD3
PSMD4
PSMD8
PSMD9
PSMD11